

29 January 1981

Madrid, Helsinki and exchanges

The reopening of the Madrid conference on the Helsinki agreements will bring very little joy to those directly concerned and little benefit to the rest of us. The proceedings so far have been one long wrangle, a measure of the extent to which East-West relations have deteriorated even since 1973, when the Final Act was signed at Helsinki. The weeks ahead are to be given over to proposals for the further elaboration of the Helsinki agreements (see page 343). For the reasons that have made the conference so far an occasion for the restatement of well-known positions, the second period is likely to see only a rehearsal of familiar opinions, known to be unacceptable to one or other of the states involved. Not much has changed since the proceedings at Madrid began. Sakharov is still in Gor'kii, for example. The replacement of President Carter by President Reagan will matter in the long run, but for the time being is likely to diminish American interest in the proceedings. The best that can be hoped for now is that the remainder of this conference will pass off without too much further damage being done.

One of the potential casualties is that provision of the Helsinki accords under which the Madrid conference itself is being held. Like many other international agreements (the Helsinki Final Act is not a treaty in the formal sense), signatories are invited to meet every so often to consider how things are working out. By comparison with Madrid, the meeting in Belgrade three years ago was more or less constructive. It accomplished little, but it did at least agree that there should be another meeting. The time is approaching in the proceedings at Madrid when people will be asking themselves whether they can properly wish on their successors in office another bout of frustration like the present. It will be unfortunate if they decide too quickly that that question should be left for decision through the normal diplomatic channels. That would assure that nothing like Madrid will happen again. That would be a misfortune.

Among international agreements, the Helsinki accords are a peculiar species. They correspond not to formal contracts among sovereign states to behave in specified ways but, rather, to the kinds of letters of intent that are signed by commercial companies which are persuaded that they have the basis for a deal but which have agreed that they may back off in certain unspecified circumstances. The Helsinki accords thus lie somewhere in between the promises to abstain from alcohol subscribed to by

members of temperance movements throughout the world and the fine print on the back of air travel tickets which binds the airline which collects the money to convey the would-be traveller to his destination provided that it is able to do so. While the Helsinki agreements were being negotiated, scepticism about their potential value abounded. In many respects, they have not worked — people's freedom to move elsewhere, for example, is still constrained. On the other hand, the remarkable freedom of journalists to visit Poland in the past few turbulent months may owe something to Helsinki. And advance notification of military manoeuvres may have done something to calm anxiety in Central Europe in recent years. The vagueness of the agreements has also, as predicted, made scope for misunderstanding but also injustice — such as the repression of the Helsinki monitoring groups in the Soviet Union. But the accords exist and, like other international agreements, could be abandoned only by further souring East-West relations.

For the scientific community, what matters most is what the accords have to say about cooperation and exchanges of people. The first flurry of excitement after Helsinki has now been overtaken by events, the banishment of Sakharov to Gor'kii only one of the more obvious. To people in the West, it is intolerable but also mystifying that scientists should be treated like this. In the East, imprisonment, exile and repression are instruments of government entirely consistent with (and thus justifiable by) the old Tsarist feudal principles that a person belongs ultimately to the state. There is nothing in the Helsinki accords to ensure that the Soviet government will behave differently towards its citizens, but there are opportunities to demonstrate that, if it did, important benefits would accrue. This is why flat declarations such as that attributed to Dr Philip Handler during the first frustrating session at Madrid can serve very little purpose. What is needed, and what Western governments should be proposing in the weeks ahead, is a much more imaginative programme of scientific exchanges than those previously arranged. Case law (and common experience) suggests that busy people do not wish to visit unfamiliar places for months on end. Why not, then, shorter visits? And why should not the governments at Madrid commit themselves to the support of participation in run-of-the-mill scientific meetings that would bring in more people from Eastern Europe, not just the Soviet Union?

Naming names in British universities

The crunch is coming for British universities. Indeed, it may have come already. The British government's decision in December that a further £30 million should be taken from the budget of the University Grants Committee in the financial year beginning in April could easily precipitate the crisis that has been looming for the past decade. For although £30 million is a mere 3.5 per cent of the budget as a whole, it will be the first time in recent memory that the universities collectively have needed to operate at less than full capacity. After three decades in which student numbers have steadily increased, universities individually are unaccustomed to turning away students for whom there is space. Worse still, the university system is not equipped to decide within itself how the consequences of even a modest reduction of business should be shared. Must all universities decline by more or

less the same proportion? Or is the misery to be concentrated in a few places and, if so, which?

Inevitably, confusion abounds. The Secretary of State for Education and Science, Mr Mark Carlisle, said on 6 January that he saw no reason why the latest budget cut should prevent universities from keeping up student numbers. The University Grants Committee, on the other hand, has issued a dark warning to universities that if they seek to compensate for reduced income from central funds by recruiting more students from the United Kingdom, they may in later years be penalized (*Nature* 22 January). The sources of these contradictory views are easily identified. The marginal cost of extra students is small and many even be zero in some universities and for some courses, so that outside observers may jump to the conclusion that less does not



mean fewer. But several universities and colleges are already so hard-pressed that they are being forced to economies that harm their teaching. Laboratory classes are especially at risk. Although the present ratio of staff to students in British universities should not be taken as a magic number, Mr Carlisle's optimism is unjustifiable. And while his statement appears to imply a promise that his own department would be willing to maintain its support for students' maintenance grants at more or less the present level, the capacity of local authorities to keep on paying tuition fees must increasingly be doubted.

The response of the University Grants Committee has been characteristically more cautious. The committee seems more keenly aware of the government's capacity (or incapacity) to pay for the maintenance and tuition of students than is the Secretary of State. Its warning that individual universities should not seek to earn their way out of trouble by recruiting more home students appears to stem from a keener sense of the British government's financial difficulties than from a judicious appraisal of the needs of the university system as a whole. In its circular to universities at the end of last year, the committee said as clearly as its habit that it will be taking steps, when it makes its financial allocations in the early summer, to control total student numbers and their distribution within universities. One consequence will no doubt be that unseemly competition for students between individual universities will be less open than it might otherwise have been. (But already many universities have partly compensated for the loss of overseas students, driven away by higher fees which were increased by a further 25 per cent last week, by taking more people from around the corner.) If the committee gets its way, there will also be no ostensible change in the present balance between the several universities in the system. No change for the time being seems to be the slogan.

There are two fallacies in this approach. First, it will not work. The competitive universities will see to that. Second, it should not be allowed to work. Further to postpone long overdue changes in the university system will merely make for greater calamity when change is eventually unavoidable. The problem is that the University Grants Committee is constitutionally unable to behave differently. Over the years, the committee has been much admired as a buffer for channelling public funds to autonomous universities in such a way that academic policy is kept out of politics. If a Member of Parliament should raise some question about the purpose of some new sociology course in some smallish university, the appropriate government spokesman will repeat the familiar declaration that academic policy is for universities, that the government habitually accepts the recommendations of the University Grants Committee on how funds should be allocated and that the inquisitive parliamentarian had better write to the vice-chancellor of the university concerned. Even though the committee is not a strictly independent public corporation — it has no bank account of its own, for example — the system worked well in the days of expansion. But the committee, which includes academics representing universities of all kinds, cannot grapple with questions of inequality — should some universities, and if so which, be treated more generously than others? Should some even be denied further support from public funds? Such questions as these are now, however, unavoidable. Who will dare to ask them?

The underlying difficulty, which should command widespread sympathy, is that there is an underlying conflict between the autonomy of the universities within the British system and the fact that they are supported largely by taxpayers' funds. Both horns of this dilemma are firmly part of British public policy. The independence of the universities is respected and, if threatened, would be fiercely fought for. Public funding, on the other hand, stems from the recognition that there is no other way in which a substantial proportion — for too long too low — of the British population should go on to higher education. The result, however, is that even — or especially — the smallest universities calculate that survival will in the long run be assured only if they ape the biggest and the best. Many of them are ill-placed to follow the successful civic universities, endowed with strong regional loyalties if not with cash, in the development of novel kinds of

courses or even local sources of funds.

The most urgent need is that the university system should work out for itself a different pattern at which to aim. All universities are equally free, but that does not imply that they are equal or that they should be dealt with as if they are. Paradoxically, the egalitarian habits of the past few decades have, if anything, served to strengthen the dominance (outside Scotland) of the universities of Oxford and Cambridge. Some other institutions (the London School of Economics and Imperial College, both parts of the University of London, among them) have kept and even strengthened their claims on British intellectual life. Some but not all of the civic universities founded roughly a century ago have branched out in distinctive ways. Some of these occasionally give Oxbridge a run for its money, both in scholarship and the competition for students. Similarly, the group of newer universities includes some with promise, perhaps because of the strength of some department or even because their location helps in the competition for students. The plain truth, however, is that many British universities, for lack of an incentive to behave differently, are undistinguished institutions with unworkable research programmes and with curricula, modelled on Oxbridge, ill-suited to the needs of their students.

The autonomy of the universities notwithstanding, the University Grants Committee is not as powerless to influence this pattern as it seems. Even as things are, a university choosing, for example, to found a medical school would quickly be told that it would have to raise the funds itself, and might even hazard support for its other expenses if it went ahead. Less expensive developments than medical schools are prudently notified to the committee in advance. (Even in these hard times, the committee seems from last month's circular creditably to be planning to earmark some of its budget for new ventures.) There is thus a *de facto* veto on major innovation. The committee is much less able to reshape the pattern of teaching if this entails that some universities should cease doing what they have always done — the committee's attempt to rationalize the teaching of agriculture in the early 1960s, only partly successful, seems to have sapped its courage. Yet present circumstances require that several universities should be asked to stop doing several familiar things.

The simplest change to engineer would be to give up the pretence that all universities are equally able to run graduate schools. This does not, of course, imply that academics need cease to be scholars — as the example of the liberal arts colleges in the United States shows vividly. What it would imply is that small departments of this or that would no longer be given tiny quotas of research studentships by the research councils, would no longer look to the University Grants Committee for the provision of a "research base" — and would no longer face the impossible task of providing a decent graduate education to handfuls of students. The same universities might be compensated by a general understanding, by the committee and others, that they should be places where four-year courses are considered normal, not exceptional. In equity, for the sake of not denying graduates from such universities a chance to make their way in research, the present delegation of the responsibility for choosing graduate students to university departments would have to be stopped. For what it is worth, the committee and the research councils between them have all the power they need to reshape the pattern of British universities in such a way.

How is the committee to summon up the courage for such a task? And how, if at all, is it to choose between the several options in the field? Its own base in the academic world is too narrow, and there is not time for a formal semi-judicial inquiry into the working of the system which it superintends. The first need is that it should decide that a radical change of pattern is needed. Then, given the autonomy of the universities, the committee needs to organize a formal and open (*i.e.* public) consultation among them. The paradox in present circumstances is that, while many academics appreciate the need for change, none of them has an interest that his or her university should make the first sacrificial move in that direction. The committee is the only agency in sight that could free this log-jam.

Johns Hopkins wins new space institute

Princeton loses space telescope ground base

Washington

After racing neck-and-neck to the finishing line, Johns Hopkins University of Baltimore has beaten Princeton University to be the site for a new research institute responsible for the operation of the space telescope to be launched early in 1985. The Space Telescope Science Institute will be managed by the Association of Universities for Research in Astronomy (AURA), a 14-university consortium. The award of the contract to AURA was announced last week by Dr Robert Frosch shortly before he left his position as administrator of the National Aeronautics and Space Administration (NASA).

AURA already operates the group of telescopes on Kitt Peak in Arizona for the federal government. It selected Johns Hopkins as the proposed site last year — even though the university itself is not a member of the consortium and does not at present have a major astronomy department. There were four other bids to operate the institute. Three of these would have put the institute at Princeton and the fourth, from the group which operates the Fermilab near Chicago, would have placed it near the particle accelerator.

The final choice was between AURA and the Princeton proposal from Associated Universities Inc., which already operates the National Radioastronomy Center and the Brookhaven National Laboratory. Dr Frosch accepted the recommendation of an evaluation team that the contract be awarded to AURA. It is thought that a deciding factor was the scope of the management scheme proposed by AURA, based largely on its Kitt Peak experience.

The new institute will be responsible for all aspects of the operation and use of the 2.4-metre optical telescope when it is launched into orbit 350 miles above the Earth from the space shuttle. Johns Hopkins University will provide a site and a new building on the edge of its Homewood campus, at a cost of about \$6 million. The AURA plan envisages an eventual staff of 150 people, including 40 full-time astronomers selected from around the world.

In addition, the institute is expected to take in up to 200 visiting scientists each year. They will be able to use the telescope 24 hours a day because, unlike surface-based optical telescopes, the sky will be permanently dark. Annual running costs are expected to be about \$10 million. Competition for the management contract has been fierce after bidding was officially

opened by NASA at the beginning of last year. There seems, however, to have been little of the political squabbling that has surrounded decisions about major scientific facilities in the past, in particular the controversies over the siting of new particle accelerators.

AURA has made much of the local facilities offered by Baltimore, which has recently been going through something of an urban renaissance. It also stressed the proximity of the city to the Goddard Space Flight Center, from which the manoeuvring of the telescope will be controlled and at which a small team from the institute will be based.

In the end, however, location is thought to have played a relatively minor part in the decision. Princeton is not much further from Goddard, and both cities have access to major international airports — one of the preconditions stipulated by NASA.

More important seems to have been a quantitative assessment made by the evaluation team of the facilities and management support being offered in each case. Johns Hopkins was not able to match the strong astronomy department offered by Princeton. Several of its physicists, however, including some members of the Applied Physics Laboratory (APL), have been closely involved in the design of the space telescope and its instrumentation.

Following NASA's decision, AURA announced that the acting director of the institute will be Professor Arthur Code, at present professor of astronomy at the University of Wisconsin in Madison. The acting deputy director will be Dr Robert Rich of APL. Dr Arthur Davidson,

professor of physics at Johns Hopkins, will be acting chief of the research support branch, and Dr William Fastie will be responsible for scientific instruments.

One factor which had concerned the AURA team while the proposals were being evaluated was the threat of federal action against Computer Sciences Corporation (CSC). The consortium had chosen the company to provide computer services, but CSC is under indictment for overcharging the federal government on time-sharing facilities.

Shortly before the management contract was awarded, however, NASA — having checked with the General Services Agency, on whose behalf the indictment has been filed — announced that although CSC has been suspended from bidding for future time-sharing contracts, it was still free to bid for technical support contracts, since these services were provided by another part of the company.

Johns Hopkins officials have welcomed the news about the institute. University president Steven Muller suggested that, given the dominant role that the space telescope is likely to play in optical observation until the end of the century, "there is true excitement at the prospect that Baltimore will now become the world capital of astronomy".

At Princeton, there was comparable disappointment. "To have hosted the institute would have been the next important step for the Princeton scientific community", one faculty member said last week. "Now we must go back and do some reassessing."

David Dickson

Dutch partners for La Palma observatory

The British Science Research Council may have found another collaborator for its optical astronomy observatory at La Palma in the Canary Islands. The Netherlands will probably join Spain, Sweden and Denmark in the collaboration to build and run the site. Last week, the council of the Netherlands Organization for the Advancement of Pure Research (ZWO) approved the deal "subject to contract", as the lawyers say. Professor Graham Smith, director of the Royal Greenwich Observatory which is coordinating the project, welcomes Dutch accession and thinks that the Netherlands may help especially with the supply of skilled manpower for management of the project.

The La Palma observatory, originally called the Northern Hemisphere Observatory, will eventually include our major British telescopes — a 1-m diameter Schmidt telescope, the 2.5-m Isaac Newton telescope (which is being transferred from Britain), a new 4.2-m telescope and a new 15-m radio telescope. Four smaller

instruments are being provided by Sweden and Denmark. Preparations for the observatory started in earnest eighteen months ago, after the protracted negotiations with Spain had been completed. Construction is well ahead, according to Professor Smith. The 1-m and the Isaac Newton telescopes are almost ready for removal to the Canary Islands and the observatory is on the verge of placing a contract for the mirrors for the 4.2-m telescope. A call has gone out to astronomers in several countries for the design of ancillary instruments for the 4.2-m telescope.

Successful bidders will be rewarded in part with observing time. The remaining contracts for the 4.2-m telescope are likely to be placed later this year, and the telescope could be working by the spring of 1985, but that date could slip by a year if the cash flow of the Science Research Council does not improve.

Much of the negotiation between the Netherlands and Britain has already been completed. In return for 20 per cent of the

cost, ZWO will be allowed to allocate 20 per cent of the observing time to Dutch astronomers. It has not yet been decided precisely how the Netherlands will help to solve the manpower shortage. When the observatory is fully operational, Professor Smith estimates that about 30 technical staff will be needed in La Palma at any one time. The difficulty is in persuading sufficient British scientists to uproot themselves for three years at a time.

ZWO has decided to pay its share out of its existing budget of about £40 million a year. Most of the money will come from the £6 million a year now spent on astronomy, implying a major shift in astronomy funding. Although most Dutch astronomers are expected to welcome the agreement, radioastronomers may feel hard done by.

The next step is for Professor van Lieshout, director of ZWO, to obtain the approval of the Dutch minister for science, who is reported to be enthusiastic. But the imminent Dutch election could mean that the issue will have to be decided by a new minister. All being well, however, Professor van Lieshout hopes that an agreement could be signed and sealed within the next four months.

Judy Redfearn

Toxic waste

Dutch dumps

Amsterdam

The cost to the Netherlands of dealing with chemical waste from years of heavy industrialization is mounting. So far about 3,000 dumps containing chemical waste have been found, 500 of them a recognized danger to public health. The Minister for Public Health and Environmental Protection estimated a few months ago that it would cost about £200 million to clear the 500 dangerous dumps, and this estimate is now £400 million.

In the village of Lekkerkerk, not far from the heavily concentrated chemical industries of Rotterdam, many buildings were found to have been built on a chemical waste dump. About 1,700 drums were recovered from the site, containing materials such as toluene and xylene from the dye industry and metals such as cadmium, zinc and lead. Some 300 houses were evacuated and 150,000 tons of polluted soil have had to be removed, while medical examinations may yet be carried out on the population. Several more chemical waste dumps have since been found in the same area.

After cases of cattle infertility and the discovery of dead birds, an investigation by the municipal environmental laboratory of Amsterdam has revealed that the Volgermeerpolder, a marshy area 5 miles from the centre of Amsterdam, contains about 10,000 drums of chemical waste from a 2,4,5-T factory previously owned by Philips-Duphar. The factory ceased

production in 1969 and was completely dismantled and dumped in the Atlantic, but its waste remained.

According to Dr H. Heida, director of the Amsterdam environmental laboratory, the drums contain a wide range of chemicals, including chlorobenzene and chlorophenol. However, some of the drums are known to contain 2,4,5-T and possibly the dioxin 2,3,7,8-tetrachlorocyclo-dibenzo-*p*-dioxin (2,3,7,8-TCDD), and it is this latter compound which it is feared may form the real danger.

Analysis of soil, water and livestock and of produce from gardens in the area has revealed the presence of chemicals from the dump in concentrations of tenths to hundredths of milligrammes per kilogramme weight.

However, the future of the dump rests on the problem of clearly identifying 2,3,7,8-TCDD. Samples from the drums have been analysed by Professor O. Hutzinger's team at Amsterdam University



The clearing of Lekkerkerk

and have also been sent to Milan for analysis by the laboratory involved in investigating the Séveso incident. A definite solution to the problem is not expected before the summer.

These pollution scandals prompted a national inventory of chemical waste in the Netherlands due to be completed by the end of 1980. So far, however, only two of the eleven Dutch provinces have reported their findings.

The problems of the Netherlands have also attracted international interest. Dr David Costle, administrator of the US Environmental Protection Agency, visited the Lekkerkerk dump while he was in the Netherlands to sign a memorandum on cooperation between the two countries. Dr Tolba of the United Nations Environmental Programme, a regular visitor to the country, is particularly interested in the chemical pollution problems.

The latest revelation was in December of last year, when small amounts of the toxic substance methyl bromide were found in drinking water in the horticultural area

between The Hague and Rotterdam. This chemical is used by market gardeners to disinfect the soil, and an estimated 2,000 tons a year are used in this area — 30 per cent of the total consumption in the European Community. The poison had apparently seeped through the PVC tubes used for private water distribution. As a result, the Minister for Public Health has made public water supply companies responsible also for private supply pipes. However, it now turns out that PVC tubes are widely used in the Netherlands for transporting water which has been tested for pollutants such as methyl bromide.

Casper Schuurin

Genetic engineering

Planning bugs

Washington

Echoes of the fierce public debates of four years ago are being heard once again in and around Cambridge, Massachusetts, as local city councils discuss the conditions they will place on the genetic engineering companies springing up in their midst.

The tones of the debate are more muted than before. And in each of the four communities — Cambridge, Waltham, Somerville and Newton — the talk this time is of negotiation rather than confrontation. Nevertheless, there is sufficient concern among the companies for the head of at least one to suggest the need for a "more coherent policy" towards local regulation, possible at the state level.

The most significantly affected so far seems to have been Genetics Institute. This is the company which has been set up by Dr Mark Ptashne, professor of molecular biology at Harvard University, and a local management consultant, Mr Tom Hexner, on privately-raised venture capital, after the Harvard faculty voted against university participation (see *Nature* 27 November 1980). Its backers include Venrock, a venture capital firm owned by the Rockefeller family, and Mr William Paley of Columbia Broadcasting.

Now the company faces a new hurdle — gaining acceptance in the local community. Last November, it applied to build a laboratory in Somerville, just inside the city's border with Cambridge and a few blocks from the university biology laboratories. The application has kindled a fierce public debate. At one point, for example, local citizens had suggested that research should be limited to P1 and P2 physical containment facilities, even stricter than restrictions in Cambridge, which allow work up to the P3 level.

City council members who were originally in favour of granting permission for the new laboratories with few strings attached have backed off in the face of the public controversy; many are preoccupied with severe budget cuts that have resulted from a recent reduction in property taxes.

At a public meeting last Thursday, the

council agreed to draw up a local ordinance setting out the conditions under which recombinant DNA research can be carried out in the city. Given the uncertainty, Genetics Institute has now withdrawn its application to build the laboratory.

Mr Hexner said last week that there would now inevitably be a delay in the company's plans. "It really depends on how long the deliberations go on and what the tenor of them is; if the tenor is favourable, then we would probably resubmit", he said. An alternative would be to look for a site in Cambridge itself. In sharp contrast with the debates of four years ago, when the then mayor Alfred Velucchi tried to ban all genetic engineering from the city, an ordinance passed in 1977 is being revised in relatively amicable circumstances.

The revisions have been precipitated by the decision of the Geneva-based Biogen Inc. to construct a new laboratory in Cambridge. Biogen's co-founder, Dr Walter Gilbert, is also a professor of molecular biology at Harvard and was a strong critic of the university's proposed involvement in Genetics Institute.

Biogen is now discussing with the local council the terms under which it will be allowed to operate its laboratory. And the Cambridge Biohazards Committee and the Cambridge Experimental Review Board are discussing a revision to the 1977 ordinance which would require, among other things, that any company or laboratory involved in recombinant DNA research should implement a medical surveillance scheme and a registry of its workers.

Such provisions have already been proposed by the Department of Labor's National Institute of Occupational Safety and Health, on the grounds that any long-term, relatively low-level effects could be traced only if adequate exposure records are kept of laboratory workers and the substances they handle. Biogen has said that it would have little difficulty in meeting such a requirement. More controversial could be the suggestion of a public debate on the acceptability of the site proposed for the laboratory.

Some members of the review board are also hoping to persuade the council to demand independent approval of the safety measures adopted by the company. This might be done by employing a consultant who would report to the city council but whose services would be partly paid for by the company.

A local ordinance including several of the proposed requirements has already been passed by the city council of Waltham, Massachusetts, home of the company Collaborative Genetics. As in Cambridge, the ordinance — which also covers research at Brandeis University — is primarily designed to ensure that all local recombinant DNA research is carried out under the safety guidelines developed by the National Institutes of Health.

In some respects, however, the ordinance is stricter than the guidelines. For example, it requires a formal medical surveillance procedure and the maintenance of work records and also prohibits the use of recombinant DNA products for trials on human patients.

With this model before them, other cities — including Newton, home of New England Nuclear — which are considering their own ordinances may find it difficult to adopt stricter rules. Mr Hexner, for example, says that Genetics Institute would be pleased to live with the Waltham ordinance and that if Somerville is too restrictive, "Waltham is there and waiting".

David Dickson

Lamont director quits

Washington

Columbia University in New York has forced the resignation of the director of the university's Lamont-Doherty Geological Observatory, Dr Mark Talwani, after complaints from senior faculty members about a lack of effective leadership.

Dr Talwani was asked by the university to resign shortly before Christmas. The university's decision was based partly on a report on the situation at the observatory, one of the world's leading centres of oceanic and seismic research, prepared by an outside committee under the chairmanship of Dr Karl K. Turekian of Yale University.

According to the university, the report praised a number of accomplishments of the observatory, but notes the existence of "serious problems" with Dr Talwani's leadership.

The committee itself did not recommend that Dr Talwani be replaced. But after further review and consultation with senior faculty members, including a meeting in October at which "almost all the senior faculty" is said by the university to have supported the search for an alternative director, Dr Talwani was asked to resign within two weeks.

Talwani himself vigorously protested at the university's decision, claiming that it had been the product of a "conspiratorial approach", and complaining to the university trustees that because no specific charges had been levelled against him, "I have had no opportunity to defend myself".

He eventually submitted his resignation as director two weeks ago, although he will remain on the university faculty as a member of its department of geological sciences. A search committee has been established to recommend a successor, and is expected to make its recommendation within a few days. Meanwhile Dr Paul G. Richards, associate director of the observatory, has been temporarily appointed its chief executive officer.

David Dickson

Soviet fishing

Whither whaling?

The Soviet Far Eastern Fishing Fleet has ceased whaling operations in the Pacific, it was announced in Vladivostok last week. The three factory ships serving the fleet are now being converted into fish-processing plants.

This is the first stage in the implementation of a pledge given at the International Whaling Commission (IWC) meeting in California in December 1978. On that occasion, the Soviet delegate, Vyacheslav Zemskii, pledged his country to phase out all whaling within five years, starting with the North Pacific operation.

The announcement that three factory ships are being converted is, however, somewhat surprising. According to a Moscow radio broadcast shortly before the July 1980 IWC meeting, the Soviet Union had only two factory ships left, the *Sovetskaya Rossiya* and the *Sovetskaya Ukrayina*, both sailing from Vladivostok. Both of these ships, the commentator stressed, carried the necessary international observers as well as Soviet marine ecologists.

A partial phasing out of the North Pacific operation began some years ago. In 1979, when the Soviet Union accepted the IWC's partial moratorium, in which fishing from factory ships would be confined to Minke whales, it had already closed down its on-shore whaling stations, and had to negotiate an exchange with the Japanese, ceding its North Pacific quota of 230 Bryde whales in return for 300 of the less valuable Minke whales from the Antarctic.

The closedown of the North Pacific operation does not, however, necessarily mean that the Soviet Union will support the idea of a total moratorium. Soviet commentators have always taken the line that such a policy is not in accordance with the ecological facts, being based on an emotionalism encouraged by some quarters for political reasons. Greenpeace observers have described the Soviet whaling flotillas as near derelict hulks which would cost many millions of rubles to refit. This suggests that the planners have realized for some time that pro-moratorium pressure might well force the Soviet whaling flotillas to disband, and that it was not therefore cost effective to keep the ships in trim.

With the recent drive for self-sufficiency in raw materials, stressed at the October 1980 Central Committee Plenum and in the Guidelines for the new Five Year Plan, what will the Soviet Union use to replace whale-oil? One possibility is the jojoba, which would probably thrive in the semi-arid steppes of Soviet Asia. Unfortunately, the jojoba plant is dioecious, and, in the natural state, male plants predominate. Jojoba culture is therefore highly labour-intensive, involving the vegetative propagation of large numbers of female

plants which then have to be bedded out with a sufficient number of males (the plant is wind pollinated.) With the problem of falling manpower, and the desire to make all agriculture more productive per man-hour, widescale introduction of the jojoba could pose considerable problems of organization.

Vera Rich

Pneumoconiosis unit

Dusty answers

Members of the staff at the British Medical Research Council's Pneumoconiosis Unit in Cardiff are alarmed about the unit's future. So too are the National Union of Mineworkers, the Agricultural Workers' Union and the Association of Scientific, Technical and Managerial Staffs, who say that they value highly the unit's work on industrial diseases. Their fears have been roused by a routine review of the unit's work, conducted by a sub-committee of the council's Physiological Systems and Disorders Board, which has apparently suggested changes at the unit. Last week, council employees and union representatives lobbied a meeting of the council at which the recommendations were to be discussed.

According to the lobbyists, the review has recommended that the unit should continue with most of its present work until a further review in 1987, but that its value should be measured increasingly by the number of contracts for research that it manages to win from industry. They also say that the review has recommended that targets be set for reductions in staff levels, that some work on allergic responses and lung function be relocated and that the unit should not receive a new high-voltage electron microscope, a replacement radiographer nor a replacement respiratory physiologist.

The subcommittee's review is a routine 3-year assessment, to which the work of all council units is subjected. Such reviews frequently lead to decisions to close some units and open others. What seems to have worried the Cardiff staff is that running down some of the work now may pave the way for closing the unit after 1987, when the present director and several other senior staff are due to retire.

Representatives of those lobbying the council's headquarters last week presented their case to the industrial liaison officer who distributed their statements among council members. Their arguments are that relocating some of the unit's work will undermine the multidisciplinary approach needed for many problems and that, by whittling away at the unit's structure now, the council may cause it to become unworkable before formal consultations on steps towards closure can be taken.

The Medical Research Council refused last week to discuss the report of the review committee or the future of the unit. It has still to inform the unit of the decision of the council meeting.

Judy Redfearn

Bulgar birthday

Two for one

"Bulgaria-1300", the unmanned space probe planned as a highlight of the country's 1300th anniversary of statehood, will now consist of two satellites. The first, Bulgaria-1300/1, will be equipped entirely with Bulgarian apparatus, and will carry out ionosphere and magnetosphere investigations. The second, Bulgaria-1300/2, will carry both Bulgarian and Soviet apparatus including several remote sensing experiments, using infrared and VHF wavelengths, and also a special Bulgarian-made multichannel camera. The experiments on Bulgaria-1300/2 will be combined with observations from a high-altitude aircraft and collated with terrestrial measurements.

No details of the experiments have yet been announced. However, according to Dr Kiril Serafimov, chairman of the Bulgarian National Committee for Space Research and Utilization, Bulgaria's "historical circumstances" are such that Bulgarian teams can take part in all five areas of space physics covered by the Comecon joint "Interkosmos" programme — space physics, meteorology, communications, biology and medicine, and remote sensing. New Bulgarian ideas, developed within Interkosmos, stated Dr Serafimov, have included Langmuir probes, ion traps and electrophotometric telescopes. During the past few years, his teams have paid particular attention to noctiluscent clouds and have reported "very interesting data" concerning discharges of auroral type over the magnetic equator. A particular Bulgarian success has been the Spektr-15 multichannel camera, whose 15 channels compare favourably with Landsat's four channels and the East German MKF-6, which has six. (In addition Spektr-15 weighs only 15.2 kg as against the MKF-6 which weighs 158 kg.) The Spektr-15 will be a major feature of Bulgaria-1300/2.

So far, however, there have been no indications of any medical or biological experiments. This seems unfortunate, as the Bulgarian space medicine experiments scheduled for the visit of their Cosmonaut Georgi Ivanov to the Salyut-6 space station were abandoned when his Soyuz transport craft failed to link up with the station. (The other "Bulgarian" parts of the programme, such as the casting of aluminium "foam" using hydrogen as dispersant, were carried out by cosmonauts Lyaknov and Ryumin, who were already on board Salyut.) Nor are there any hints that Soviet birthday presents to Bulgaria will respond to the broad hints that the Bulgarian space planners have been dropping ever since Ivanov's mission was aborted — another, and this time more successful, flight of a Bulgarian cosmonaut. Other states have done better, but are not as old. Vera Rich

Reef aggregation

A new society and a new journal concerned with research on coral reefs have sprung from the frustrations of a group of interested scientists at the difficulties of funding this interdisciplinary work. The new society, called the International Society of Reef Studies, was formed after a meeting held at Churchill College, Cambridge last month.

The first president of the society is Dr D.R. Stoddart from the Department of Geography at the University of Cambridge, who was the leader of the Royal Society's expedition to the Indian Ocean island of Aldabra in the 1960s. The council of the society includes members from mainland Europe, North America, Australasia and the Far East.

One of the objectives of the society — to provide a forum of coral reef research — has been met by the proposed publication by Springer-Verlag of a quarterly journal, *Coral Reefs*, due to appear next year. The new journal will be more concerned with the structure and dynamics of coral reefs than with the flora and fauna, and will thus be complementary to the *Atoll Research Bulletin*, published by the Smithsonian Institution.

The society also plans to take over the management of the international symposia on the subject, of which the fourth is due to be held in Manila in May. Its formation illustrates the difficulties of recruiting funds for the support of research in fields that do not fall squarely within the terms of reference of established grant-making agencies, which have been acutely felt in Britain in recent years. Feeling has run high that governments are interested in coral reefs only in crises.

Brazilian ecology

Polluting by decree

São Paulo

Economic pressures have caused the Brazilian government to postpone legislation banning the use of non-biodegradable detergents which was due to come into force in January 1981.

A decree signed by President Figueiredo last month puts back the starting date for the new law by two years. Industrial pollution is now a serious problem in Brazil. In the state of São Paulo all the rivers are biologically dead and the appearance of foam on rivers is now considered normal, according to the head of the Environmental Technology Agency. Several waste-treatment plants due to open in 1981 in São Paulo will have to be abandoned if the foaming persists.

Although the intended ban on non-biodegradable detergents was drafted four years ago, nothing was done to make it

possible to comply with the deadline. Manufacture of the non-biodegradable component of detergents, alquilbenzene sulphonate (ABS), is a monopoly of the Atlantic Richfield subsidiary Empresa Carioca de Detergentes (EMCA). Its production at 45,000 tons a year covers all Brazil's needs. SPUMA, the only company importing biodegradable linear alquilbenzene (LAB) was forced to close down more than a year ago because of prohibitive transport costs from its factory in Manaus in the Amazon region to the main consumption centres of São Paulo and Rio 4,000 miles away. The government chose not to help SPUMA by reducing import taxes and the company's machinery was bought up by a multinational company and left idle.

The Association of Cleaning Products Industries has been accused of lobbying against the entry of biodegradable detergents on to the Brazilian market. The association pleaded for a four-year delay in passing the new law on the grounds that the cost of importing the annual requirement of 35,000 tons of LAB would amount to \$35 million in hard currency, increasing the \$330 million deficit in the Brazilian balance of payments.

What this ignores, however, is the fact that the Brazilian owned company DETEN is now building a plant for the manufacture of LAB in the huge petrochemical complex under construction in Salvador, Bahia. Production is to start in April, reaching 25,000 tons by the end of 1981 and satisfying the expected Brazilian requirement of 70,000 tons per year by 1983. By then the multinational-linked EMCA will also be producing LAB in Brazil at the rate of 25,000 tons per year.

The economic argument for keeping the Brazilian market open to non-biodegradable detergents is further undermined by the statement from the Health Commission of the Brazilian Congress that EMCA had received a firm offer from CONOCO, a Shell subsidiary, to buy up all of EMCA's yearly production of ABS for \$36 million for export to other even less ecologically minded countries.

When recommending the President to postpone the ban, the Minister of Industry and Commerce ignored the ecological arguments and cited only the views of the Association of Cleaning Products Industries. The Health Minister in his turn agreed to the postponement on the grounds that health issues were not involved.

The new decree leaves it up to the Ministry of Health to set "indices of biodegradability" over the next two years compatible with the proportions of LAB and ABS on the market. But the Minister of Industry has stated that the responsibility for all future regulations concerning the production, import and distribution of biodegradable products will remain with the executive branch and not with the legislature.

Maurice Bazin

Security conference

Back to Madrid

The Madrid review conference of the Helsinki Final Act reconvened on Tuesday this week for a further two months. Attention now turns from the presentation of reports on progress by the signatory states to the potentially even more acrimonious question of future prospects.

Since the previous review conference in Belgrade in 1977, East-West relations have cooled. One consequence at Madrid was the long procedural wrangle before the conference could begin, partly over whether Afghanistan should be on the agenda.

Later meetings (which were closed to the public) saw much plain speaking, notably from Dr Philip Handler, president of the US National Academy of Sciences. He pointed out that the Hamburg "scientific forum" last February had been devoid of meaning because of Western scientists' concern about "serious infringements of the human rights and freedoms of too many of their colleagues in the East". He concluded with a quotation from Sakharov: "Intellectual freedom is essential to human society — freedom to obtain and distribute information, freedom for open-minded and fearless debate, and freedom from pressure by officialdom and prejudice".

Intellectual freedom concerns only the third of the three "baskets" into which the Helsinki process is conventionally divided. The term "basket" is a relic of the Helsinki Conference in 1973 during which it was decided that there should be three large filing baskets into which delegates could drop their suggestions on, respectively, military matters, trade/technology exchange and human rights.

Western signatories have always maintained, however, that this purely practical procedure did not change the essentially indivisible nature of the process of détente. The socialist bloc, however, would prefer to deal with the three "baskets" in isolation, which, in practice, means concentrating on the military aspects. Indeed, it has often suggested that the Western emphasis on human rights is itself a breach of the Final Act, being an interference in the internal affairs of signatory states.

The discussion of human rights at Madrid has been further complicated by semantics. The United Nations has made a working distinction between "legal" rights and "programme" rights, such as the right to an adequate living standard and the highest attainable standards of physical and mental health. Western signatories insist on the narrower meaning, Soviet spokesmen on the broader.

Proposals tabled for discussion in the next two months are of two kinds, concrete and declaratory. The French, for example, will urge that military "confidence building measures" should be extended to cover the whole of Europe as far as the

Indians home from home

Lucknow

In an attempt to recoup some of the talent lost through its relentless "brain-drain", the Indian government has reached an agreement with the United Nations Development Programme (UNDP) to implement a project known as TOKTEN, or the "Transfer of Know-how Through Expatriate Nationals".

Under the project, which it is hoped will further India's National Development Plan, Indian expatriates who are specialists in various technical and sociological fields will be invited to return to India for periods ranging from one week to three months. During their stay contributors would be expected to pass on their skills, act as consultants in their specialist field and make suggestions on issues of policy. They would help put their ideas into practice while in India and continue to give advice afterwards.

Responsibility for running the project will lie with a committee specially formed by the government's Council of Scientific and Industrial Research. Its members include representatives from ministries, government departments, public and private sector organizations, the University Grants Committee and the Indian branch of UNDP. Meanwhile UNDP has allocated an initial \$100,000 for travel and living expenses.

So far about 300 Indians — including doctors, engineers, scientists and social scientists — have agreed to take part in the TOKTEN project. **Zaka Imam**

Urals. At present, notification is required only within a frontier zone of 250 km.

The Soviets, for their part, have called for a European disarmament conference. Poland has proposed Warsaw as a suitable venue, and October 1981 as a possible date. A recent proposal by President Urho Kekkonen of Finland, which is likely to be raised again at Madrid, is known to have strong Soviet backing: the establishment of a nuclear-free zone in Northern Europe. (They are incensed, however, at Danish suggestions that "some regions of the USSR" should be included in the zone.)

Basket Two proposals (trade and technology transfer) will have to be discussed against the background of official trade sanctions imposed on the Soviet Union following its intervention in Afghanistan, and of "personal boycott" pledges by individual scientists and the United States academy's moratorium on exchanges. Major advances are unlikely.

Basket Three will carry a whole range of proposals, from the easing of visa restrictions and the "reunification of families" (the clause of the Final Act most frequently invoked by Soviet Jews wishing to emigrate) to the rights of private citizens to monitor the implementation of the Helsinki accords. **Vera Rich**

CORRESPONDENCE

Galileo's trial

SIR — Your recent leading article "Why bother to rehabilitate Galileo?" (*Nature* 30 October 1980; 287, 767–768) begins with a pertinent review of Galileo's trial. The last part of the article unfortunately becomes an arbitrary critique of the twentieth century Catholic Church which does not belong in a scientific journal like *Nature*.

More important though, some errors are present. For instance, it is stated that directed evolution plays a "big part in the Church's view of biology", that the Big Bang is "leaped at" by the Church and that the scientific failure of these two theories would place the Church in an indefensible position. These three statements are invalid.

The official Church never taught "directed evolution". Teilhard de Chardin never was a "doctor of the Church" and the official theology of the Catholic Church is still founded on the works of men such as St Thomas Aquinas, who admitted the possibility that the world apparently acts as moved by the laws of fate and necessity (see also St Augustine).

The Big Bang was once kindly considered (but never affirmed) by Pope Pius XII. And some modern astronomers actually claim that the Big Bang is a proof of God (which is either stupid or blasphemous, and perhaps both). The finiteness or infiniteness of space, time and the number of intelligent beings in the Universe are not fundamental problems and have no bearing on the primary question "Why existence?". Science and God, for the Catholic Church, do leave the individual free to answer such questions for himself, and it seems useful that you should be aware of the affirmation of Pius XII, Paul VI and John-Paul II that science and faith each have their own field.

The article "Why bother to rehabilitate Galileo?" ends with categorical criticism of the Church's moral (sexual) recommendations. What is the point of this? The Church proposes a way of life (responsible paternity) to people who believe in the Church and in the eternal destiny of men. Ethics is not a consequence of science.

Then you continue to say peremptorily that "The Church's views on birth control . . . are a means of making millions of small-time martyrs, pathetic contemporary analogues of Galileo" — what can be said of abortion, which involves an act similar to that of the men who condemned Galileo?

HENRI REBOUL

Université des Sciences et
Techniques du Languedoc,
Montpellier, France

Don't be vague

SIR — Mr Ronald Reagan may indeed be conservative but he is surely not so backward-looking as to choose a British First World War general as his Secretary of State (*Nature* 18/25 December 1980, p.631).

DAVID PIKE

King's College Hospital,
London SE5, UK

Secretary of State Alexander Haig was
inexorably confused with Earl (Douglas)
Haig (1861–1928) — Ed. *Nature*

Evolutionary thought

SIR — At a time when irrationalism is gaining ground against science in popular culture I feel that I must echo some of Dr Halstead's misgivings about the use of cladistics to explain the process of evolution¹. However, his castigation of Engels as the *eminence grise* behind cladistics is unfounded and is based upon a misunderstanding of how dialectical materialism explains discontinuous development in both nature and society.

Halstead (with Popper) sees Marxism as "working to spark off the revolution"². This is the sort of ultra-leftism espoused by the early German Communist Party and later by Stalin in his "Third Period". Lenin in "Left-wing Communism — an infantile disorder" answered the former and Trotsky the latter in "Revolution Betrayed". In the biological sciences, organic change may be either gradual or discontinuous. Darwin by over-reacting to "leaps" in development, which he saw as the "new creations" of his religious opponents, ended up by formulating a theory of inheritance based on the blending of acquired characters — pangenesis. This ultra-gradualism was overthrown by genetics which reinstated the view of saltatory leaps, now based on mutations. Dialectical materialism properly understood proposes neither an opening for deism nor mutational leaps. It does propose the transformation of quantitative developments (genetic mutations and changes in ecology) into a qualitative leap in development (not in structure but in selection pressure). A genetic mutation *per se* is a leap, but not a dialectical leap.

The clearest examples of dialectical change in evolution are in the area of so-called macroevolution, where quantitative genetic changes taking place to fit an organism for one habitat pre-adapt the organism to "leap" into another ecological niche where the selection pressure is qualitatively different. Darwin describes this type of development in the speciation of the Galápagos finches, although not referring to it as a leap. The early diversification of the mammals has been described by Olson³ as due to "Co-ordination and integration of such modifications (as) brought some members of a particular radiation to the threshold of a new radiation". Similarly, many workers explain the initiation of major periods of diversification or the invasion of new environments by such a "quantitative to qualitative and back to quantitative" hypothesis.

Even G. G. Simpson, to whom the mechanistic view of cladism may or may not have been anathema, was by no means averse to the idea of leaps. I quote: "Even within the staid horse family, which seems as a whole to be progressing rather steadily through the Tertiary, close examination shows the rates varying considerably. More broadly, evolution commonly seems to proceed in spurts and pauses in an apparently erratic way"⁴.

MIKE HOWGATE

Department of Zoology,
University College London,
London WC1, UK

1. Halstead L. B. *Nature* 288, 208 (1980).
2. Halstead L. B. *New Scientist*, 17 July 1980, p.215.
3. Olson E. C. *Vertebrate Paleozoology*, 673 (Wiley-Interscience, 1971).
4. Simpson G. G. *The Meaning of Evolution*, 103 (Yale University Press, 1976).

Engels on Darwin

SIR — How right Dr Halstead is to remind us of the lengths to which Marxists have gone, and will no doubt continue to go, in their efforts to place their materialist theory of history on a scientific par with Darwin's evolutionary theory. Engels, no less, argued, in his oration at Marx's funeral, that "... just as Darwin discovered the law of development of organic nature, so Marx discovered the law of development of human history. . . ."

Incidentally, it is worth recalling that the only Englishman at the graveside to hear this, was Marx's erstwhile young friend Ray Lankester FRS. Lankester was later knighted and served as the Director of the Natural History Departments of the British Museum (1898–1907), and is still remembered for his reclassification of the Talpidae collection. The plot thickens! HARRY CRABTREE
Birmingham 13, UK

Motorbike safety

SIR — A recent letter to *Nature*¹ concerns the observation that an object appears to maintain constant size as its distance from the observer changes. It is simply "seen" as a full-sized object at varying distance. This size constancy fails quite suddenly when the object subtends an angle at the observer of less than half a degree; it is then seen as a small object, presumably of less immediate importance.

This phenomenon could be a contributory factor in a fairly common type of road traffic accident, that caused when a driver pulls out unexpectedly in front of motorcyclist who is then unable to stop.

An approaching motorcyclist presents an irregular silhouette subtending an average angle of less than half a degree at any distance greater than about 200 feet. If size constancy fails below half a degree, the motorcyclist would only be seen as small, and would therefore be less likely to register subconsciously as a real full-sized danger, until within about 200 feet.

The shortest possible stopping distances for a moving vehicle under good conditions are reported in the highway code as follows

Vehicle speed (m.p.h.)	30	40	50	60	70
Stopping distance (ft)	75	120	175	240	315

Interpolation suggests that when a motorcycle is approaching at over 54 m.p.h., so that its stopping distance is greater than the critical 200 feet, an accident might be inevitable, particularly if either driver was not concentrating on "seeing" the other vehicle at the crucial time. Many similar calculations are possible. A convex rear view mirror reduces the image size and so can make it possible for larger vehicles to approach within their stopping distances before they, too, subtend the critical angle of half a degree.

Much work has been done on the effect of factors such as brightness, contrast, colour, shape, size (and styling features disguising shape and size) on the visibility of an object. The results of this work might valuably be applied to problems of the type outlined above

G. K. MCGINTY

Redhill, Surrey, UK

1. Ross J., Jenkins B. & Johnstone J. R. *Nature* 283, 473–474 (1980).

NEWS AND VIEWS

The Nobel awards for medicine or physiology

from Sir Peter Medawar CH, FRS

THE joint award of the 1980 Nobel prize for medicine or physiology to Jean Dausset, George Snell and Baruj Benacerraf has been well received by immunologists and medical biologists generally. The general public and the scientific press also found the award easy to understand: Jean Dausset was prime mover in the discovery in human beings of an important system — 'HLA' — of tissue groups that affect the outcome of organ transplantation in a way analogous to that in which the principal blood groups — ABO — affect the outcome of blood transfusions.

Its relevance to tissue transplantation is not, however, by any means the most important aspect of the discovery of the so-called 'major histocompatibility complex' (MHC) of human beings and of the analogous — perhaps homologous — MHC of mice, known as H-2.

The great and abiding importance of the HLA system is twofold. First, it defines a system of polymorphism — of stable genetic differentiation — in human beings which has uncovered, as no other procedure could have done, the genetic basis of differences in susceptibility to multiple sclerosis, ankylosing spondylitis and the juvenile (insulin-dependent) form of diabetes. Even more important, it has become increasingly apparent that the MHC of man as of the mouse defines a whole system of cell-surface markers which are of the utmost importance for the organism's power to discriminate between self and non-self; indeed, an organism's power to react to a number of immunogens as antigens seems to depend on their being signalled by the gene products of the MHC.

Although Jean Dausset was prime mover in the recognition and definition of the HLA complex, this discovery — as with all others of comparable stature — was the work of a number of medical scientists standing shoulder to shoulder: the names that come instantly to mind are those of Bernard Amos, Richard Batchelor, Walter Bodmer, Ruggero Cepellini, Rose Payne, Jan van Rood and Paul Terasaki — a splendid example of international co-operation promoted by the Transplantation Society's generosity and good sense in organizing histocompatibility workshops at which all the principals met to exchange ideas, knowhow and reagents.

To explain George Snell's award we must go back to the early days of tumour transplantation. In the first decade of this century it was widely believed that a study of the factors promoting or retarding the growth of transplanted tumours held promise of a prevention or cure of malignant growth; but by the end of the decade, thanks mainly to the criticism of Peyton Rous, it came to be recognized that the study of tumour transplantation was a study of transplantation rather than of tumours, for in the days before homozygous strains of mice became available, tumour transplants were all of the kind known as 'allografts' and as such aroused immunological rejection reactions.

A very great deal of the work was done with what might be called 'any old mice'; it made no kind of sense and — what is worse — still cannot be interpreted today. This judgement applies with especial force to work on 'immunization' against transplanted tumours by the previous inoculation into their intended recipients of tumour cells of the same kind or of a variety of normal tissues.

Matters were in a desperate state until they were taken in hand by Little, Strong, Bittner, Johnson, Bagg and George Snell of the Roscoe B. Jackson Memorial Laboratory — the men who originated the inbred strains used throughout the world today. Using genetically uniform mice it became possible to formulate the elementary genetic ground rules of transplantation of normal tissues or of tumours. George Snell went further and worked out in great detail how the outcome of transplantations is governed by the genes that comprise the MHC of mice (H-2), first serologically recognized by Peter Gorer of Guy's Hospital, London, who if he had lived, would certainly have shared the Nobel prize. The study of H-2 provided the technology and the conceptual background that made possible the recognition of the MHC of man.

This has been the international year of the MHC, for the award to Baruj Benacerraf also relates to the MHC — in particular with that sub-region of it that has to do with 'Ir' (short for immune response) genes. When Benacerraf and H.O. McDavitt first recognized this sub-region it was widely thought that they had discovered the gene

that coded for the antigen-specific T-cell receptor, but it now seems more likely that the Ir region is involved with the presentation of antigen to T-helper cells.

The work of this year's recipients of the Nobel award for medicine or physiology has a twofold lesson to teach. The first is of the immense fruitfulness of collaboration between mice and men in the solution of medical biological problems. The second, not quite so obvious, is that the discovery of HLA provides an exemplary case history in rebuttal of the notion that scientific discovery can be premeditated. What advice, for example, should one give to a young man intent on finding out the genetic basis of differences in susceptibility to multiple sclerosis, ankylosing spondylitis or insulin-dependent diabetes? It would be the *reductio ad absurdum* of the notion that scientific discovery can be premeditated to recommend that the young research worker should embark on a study of tumour transplantation, specifying the use of heterozygous mice, until its contradictions and vicissitudes drove him nearly mad whereupon he should proceed at once to work out its genetic and serological basis, from which the results could be translated to human beings and the rest would follow.

In days when businessmen and others who ought to know better are advocating a diminution or even abandonment of 'pure' research in favour of its practical applications it cannot be too strongly emphasized that discoveries such as HLA — as of diagnostic X-ray radiography — were made by the ordinary processes of scientific discovery, slow, messy and uneconomic though they undoubtedly are. It is in fact only unworldly impractical daydreamers who think that scientific discoveries can be premeditated or can be arrived at by the working of some calculus, some organon of discovery. It was the tough practically minded, no-nonsense men of affairs — particularly of great affairs — the world's Rockefeller, Sloans, Nuffields and Wolfsons who proved to understand most clearly the need and to provide most generously the means for what is so mistakenly called 'pure' research. □

Sir Peter Medawar, CH, FRS is Head of the Transplantation Biology Section of the Medical Research Council, UK.

Meteorite breakup

from Michael R. Dence

INTEREST has revived in crater groupings on the solid planets since the report that some asteroids may be double (Tedesco *Science* **203**, 905; 1979). Earlier debates between those who considered that groups resulted from two or more simultaneous impacts and those who considered them to be chance clusters of unconnected events have run in favour of chance (see Woronow *Icarus* **34**, 324; 1978). Yet some crater associations, most notably on Earth, seem to require coeval impact events. Can all such groups be accounted for by breakup of single bodies before impact or are they evidence for meteoroid pairs in space? To answer this question statistical studies have now been augmented by attempts to derive a quantitative understanding of meteorite disruption by an atmosphere or by tidal rupture.

The breakup of meteoroids or of re-entering satellites has been described by many eyewitnesses. Disintegration of meteorites has also been photographed and led in the case of the Innisfree Alberta fall of February 1977 to the recovery of nine fragments of the aerolite (Halliday, Blackwell & Griffin *J. R. astr. Soc. Can.* **72**, 15; 1978). Much larger meteorite showers may contain many hundreds of fragments. In such falls the meteorite fragments, or the craters they produce if they land with sufficient velocity, are usually distributed over an area elongated in the direction of flight of the bolide. The strewn field or scatter ellipse is commonly only a few square kilometres in area, but, as in the case of the great 8 March 1976 fall at Kirin, North China, may cover several hundred square kilometres. Although, on average, the scatter ellipse is one-third as wide as it is long, some are almost circular while others show a down-range dispersion more than eight times the cross-range scatter. Commonly, but by no means invariably, the largest fragments or craters are found at the down-range end of the ellipse.

What causes this scatter? The general principles have been understood for some time and a few falls have been analysed in considerable detail. The veteran Russian meteoriticist E. L. Krinov concluded that the massive Sikhote-Alin fall of 12 February 1947 came from an iron meteorite weighing more than 70 tonnes which broke up in three stages. From the pattern of secondary scatter ellipses superimposed on the main ellipse, Krinov deduced that there was an initial high-altitude breakup followed by two additional disruptions, the last at an altitude of 6 km or less (*Meteoritics* **9**, 255; 1974).

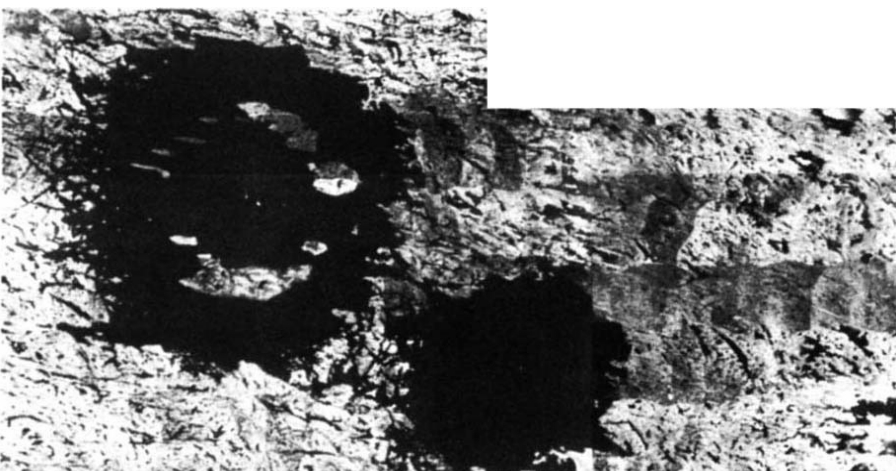
This and other large falls, mostly of iron meteorites, have now been modelled quantitatively by Passey and Melosh (*Icarus* **42**, 211; 1980) with some unexpected results. They trace the history of individual meteoroids of varying size and crushing strength after they enter the atmosphere at an altitude of 70 km. Velocity and entry angle are also varied in the calculations. Initial fragmentation is taken to be largely the result of aerodynamic pressures on pre-existing defects or planes of weakness, though they recognize that thermomechanical stresses may also be effective. The mechanics of breakup are not fully understood and the estimation of appropriate properties such as crushing strengths is one of the main uncertainties of the analysis. Following breakup, dispersion down-range is considered to result mainly from atmospheric drag. Since the small fragments are decelerated more effectively than the large, they tend to fall short of the main mass. Lift may counter the effect of drag, but Passey and Melosh argue that it is only effective for fragments lighter than 10^2 kg. For larger masses they conclude that lift is a small fraction ($\sim 10^{-3}$) of drag, otherwise many more would lie down-range of the main mass, or crater, than is generally observed. Parenthetically it may be noted that Sikhote-Alin is one instance

where many sizeable craters are down-range from the largest.

A feature of their analysis is the attention Passey and Melosh give to cross-range dispersion. This they attribute to a combination of the effects of bow shock interaction, spin of the meteoroids acquired before entry, and spreading induced by crushing as the meteoroid decelerates. All of these effects they judge to be quantitatively similar and capable of explaining most of the observed scatter transverse to the direction of flight. They calculate that maximum separation is achieved when breakup occurs at an altitude of about 15 km.

More of a problem is their conclusion that, for meteoroid masses of 10^4 to 10^9 kg, multiple crater fields result when the angle of entry of a bolide is less than 30° . Indeed, in analysing specific terrestrial crater fields, formed for the most part by iron meteorites, they generally favour entry angles of less than 20° . However, the examples they draw on include 70 per cent of the terrestrial inventory of craters 20 to 200 m across with associated meteorite fragments. As the probability of the entry angle being 30° is only one in four and is even less for lower angles, it is evident that the result implies a marked deficiency of craters from high-angle impacts. As there are no other grounds for believing this to be the case, their analysis requires further examination. One method of expanding the set of conditions for producing crater fields is to lower the strength assumed for the meteoroids. Passey and Melosh adopt yield strengths between 10^7 and 5×10^8 Nm^{-2} in their representative calculations. This range may be appropriate for the low-altitude, final breakup of the Sikhote-Alin meteoroid, but its initial, high-altitude disruption implies considerably lower strength for the main mass. We are brought again to the need for an improved theory of meteorite fragmentation.

Further problems arise when bodies of mass greater than 10^9 kg are considered. Basically, much larger bodies on entering the atmosphere may rupture but no combination of atmospheric deceleration, lift, bow shock interactions or gravity will allow sufficient separation for the formation of more than one crater. For typical velocities of 15 to 20 km s^{-1} , separations of the order of 20 times the radius of the impacting bodies are required to achieve clear separation of the resulting craters. Nonetheless, pairs of large craters, formed by bodies of the order of 10^{12} kg, are found on Earth. The best known are the 32 km and 24 km diameter pair at Lac à l'Eau Claire (Clearwater Lakes, see figure), northern Quebec, Canada and the 24 km Rieskessel and 3.5 km Steinheim Basin combination in southern Germany. The former have a centre-to-centre separation of 31 km, the latter of 45 km. Radiometric



Michael R. Dence is Acting Director of the Gravity and Geodynamics Division of Energy, Mines and Resources, Ottawa.

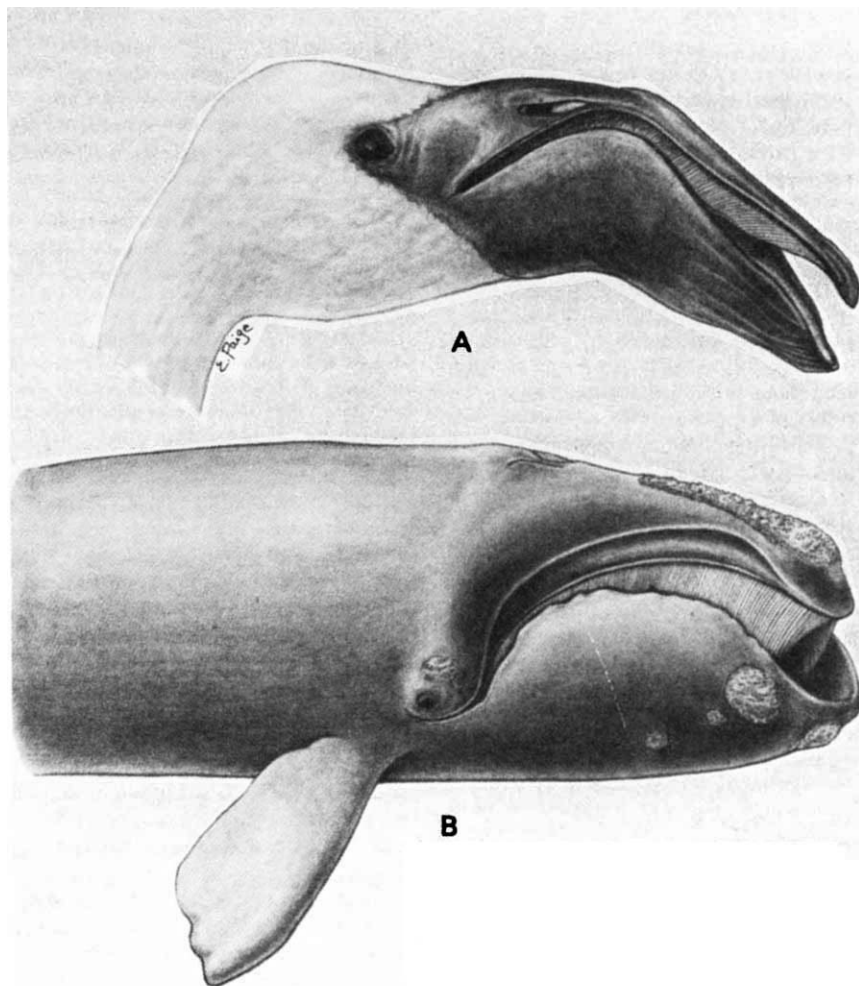
Flamingos, stilts and whales

from Andrew Milner

THE phylogenetic position of the flamingos (family Phoenicopteridae) has long been the subject of controversy among avian systematists with some workers supporting relationship with the Anseriformes (ducks and geese) whilst others favour association with the Ciconiiformes (storks, herons and ibis).

The evidence has been reviewed recently by Olson and Feduccia (*Relationships and Evolution of Flamingos (Aves: Phoenicopteridae)* Smithsonian Contributions to Zoology **316**, Smithsonian Institution Press; 1980) who have developed an earlier proposal by Feduccia that the flamingos fall within the Charadriiformes (waders, gulls and auks). Characteristics of the musculature, the skeleton, the natal down and the endoparasitic cestodes all lead to the conclusion that the closest living relatives of the flamingos are the Recurvirostridae (stilts and avocets) and, more specifically, the Australian Banded Stilt (*Cladorhynchus*). This remarkable bird lives in colonies frequenting temporary salt lakes in southern Australia and its behaviour, pattern of breeding and life history all bear specific similarities to those of flamingos. The earliest certain fossil flamingo from the Eocene of Wyoming is described by Olson and Feduccia and supports their hypothesis by being morphologically intermediate between recurvirostrids and flamingos. One phylogenetic anomaly, the presence of anseriform-type feather lice in flamingos but not in *Cladorhynchus*, is attributed to the presumed common lifestyle and physical proximity of early flamingos and the early colonial Anseriformes such as *Preshyornis*.

The filter-feeding mechanism of flamingos is structurally and mechanically different from that of ducks but is remarkably convergent with that of the baleen whales, particularly the Right



Whale! Despite the differences in scale and orientation during feeding, the two filter feeders share structures such as a large fleshy tongue accommodated in a deep-sided lower jaw and a recurved rostrum permitting the presence of a row of elongate filtering devices suspended from each upper jaw. The similarity in structure and function of the feeding apparatus of whale and wader suggests to

the authors that the 'crooked beak' of the flamingo may have been a fundamental adaptation to filter-feeding which preceded the 'inverted-head' feeding posture rather than following it as has generally been assumed.

Andrew R. Milner is in the Department of Zoology, Birkbeck College, University of London.

and geological dating leaves little doubt that each pair was formed simultaneously. In addition, in neither is there a structural basis for assuming a highly oblique impact angle. It seems that both pairs may indeed be the product of the impact of double meteorites.

The idea is in fact not new having been examined by R. W. Tanner in 1963 (*J. R. astr. Soc. Can.* **57**, 109). Recently it has been considered in greater detail by P. D. Noerdlinger (*Icarus* in the press) who, after an exhaustive study of tidal disruption mechanics, confirms that Tanner's analysis is essentially correct. In the general case of tidal rupture at a distance of several Earth radii the separation that fragments achieve approximates only the initial radius

of the parent body. Overlapping or nested craters will result. It is conceivable that the separation found at the terrestrial crater pairs might be achieved by tidal or atmospheric rupture of a body of low strength which, entering at an oblique angle, became a satellite for several orbits. Clearly such an event has an extremely low probability, but the Ries-Steinheim and Clearwater pairs are two occurrences in a total population of approximately two dozen large (~ 20 km) craters (Grieve & Robertson *Icarus* **38**, 212; 1979). Once again the existing crater statistics argue that multiple cratering events, even in this size range, are not of great rarity.

On the other hand Noerdlinger contends that the population of meteoroid doubles is

also probably small. Rupturing an asteroid so that the pieces remain in close association requires a delicate balance between the energy of fragmentation, gravity and angular momentum. The association may also be difficult to maintain for any appreciable time. The entire life cycle of double meteoroids and asteroids is in need of a careful, quantitative evaluation of the circumstance under which they can form and their subsequent stability before final capture, separation or disintegration. Such a study would be relevant to another important issue — the formation and lifetime of ring systems as now revealed in spectacular detail by the Voyager images of Jupiter and Saturn.

Cement — a respectable material?

from D.D. Double

FAMILIARITY breeds contempt. We see cement everywhere, mixed with sand to form mortar between bricks, mixed with pebble aggregate to make concrete. World-wide production of cement approaches a billion (10^9) tonnes per annum and of this 95 per cent is Portland cement, the remainder being more specialized products such as high alumina cement. This figure very substantially exceeds the production of steel, traditionally its nearest counterpart in large-scale engineering construction. Yet, in spite of its obvious technological importance, relatively little has been done to explore the possibilities of cement other than as a convenient gluing agent in a cheap bulk filler material. Relegated to this rather unsophisticated role, outside the established industry, cement has not received the attention that it deserves.

But this situation may be changing. In general, there is increasing interest in inorganic materials that can be produced cheaply and which might compete with or even replace other common materials which are more costly in economic or ecological terms^{1,2}. In this context, silicate- and aluminate-based systems seem an attractive proposition and one of the obvious sources for development is cement (principally Portland but also other hydraulic cements). Materials scientists, who have hitherto mostly ignored cement as a not-quite-respectable material, are now taking a different view.

Portland cement is remarkably cheap and materials required for its manufacture (limestone or chalk, and clay) are widely abundant. In terms of the energy costs for production it compares very favourably with other materials.

Energy cost for production (per unit volume, relative to cost of cement)

Portland cement	1
Glass	1.3
Aluminium	35
Steel	20
Polystyrene	5

Whereas other materials, such as glasses, ceramics and metals, usually require high-temperature, and therefore expensive, shape-forming processes, cement develops its strength by reaction with water at ambient temperatures. Before setting, it is easily castable and made into complex shapes. With a density around 2.4, compared for example with that of aluminium 2.7, a hardened cement paste is relatively light. It has low thermal and electrical conductivity, is inherently non-combustible and, as far as one can tell, is generally non-toxic. Young's modulus has quite a respectable value around 20 GN m^{-2} .

With all these potential advantages, one might wonder why cement has not been used more widely in more diverse ways. The reason is quite simple — as normally

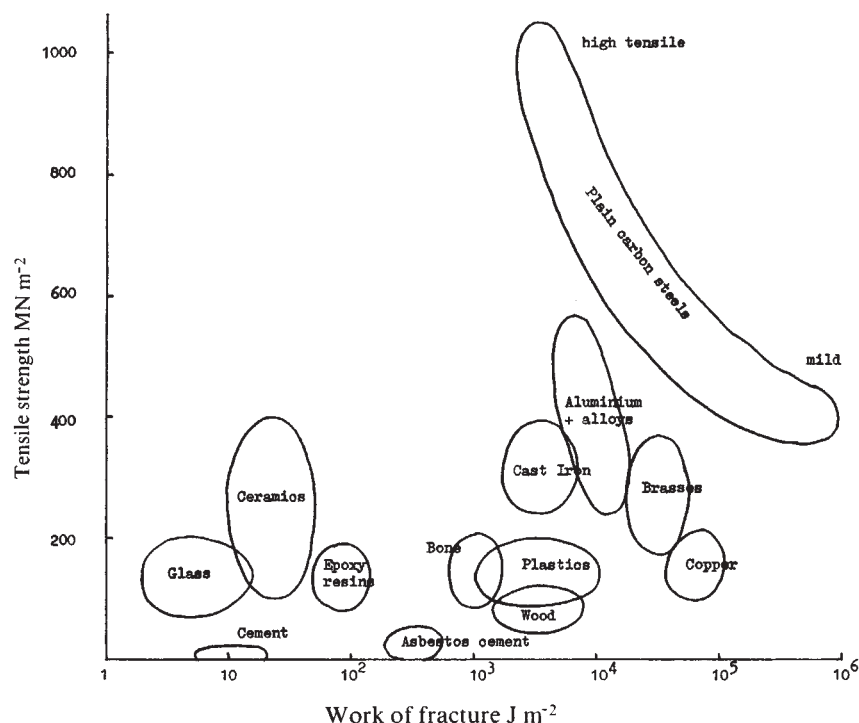
made, it is very weak and brittle. Although a hardened cement can sustain quite appreciable compressive loads, its tensile strength and fracture toughness — two parameters that often determine the practical use of a material — are disappointingly low, respectively 5–10 MN m^{-2} and about 20 J m^{-2} . The situation in relation to other common materials is illustrated in figure 1. Engineers are well aware of the mechanical deficiencies of cement as a binder. Wide safety margins are usually used when dealing with structural concrete sections and methods are devised (reinforcing, prestressing, cantilevering) to avoid putting concrete under any substantial tensile stress.

It is obvious that if cement is to be exploited in new and unconventional ways then improvements have to be made in the mechanical properties. Cement will never be as strong as steel but, referring to figure 1, one can see that only relatively modest improvements are required to bring it on a par with a range of other materials. Cement could become a serious contender at a tensile strength of 100 MN m^{-2} and 10^3 – 10^4 J m^{-2} in a work of fracture. The question is whether such improvements are possible.

How does nature deal with the problem of forming strong inorganic solids in an aqueous medium at normal temperatures? The classic example is the calcium carbonate found in shells of marine organisms. Measurements of mother-of-pearl shell (nacre) show that it has a modulus around 100 GN m^{-2} , a tensile strength between 70 and 100 MN m^{-2} and a fracture energy greater than 1,500 J m^{-2} (ref. 3). It is clearly a much more impressive material than $CaCO_3$ in the form of chalk, limestone or even the best quality marble. It achieves these properties by developing an intricate structure of low porosity consisting of plates of crystalline $CaCO_3$ (aragonite) held together by very thin layers of a protein polymer 'glue' (figure 2). Limpet shells have a more complex cross-laminated structure and certain crab shells show an interesting, and presumably purposeful alternation between calcitic and aragonitic crystal forms of $CaCO_3$. These examples demonstrate that, through appropriate reorganization of the micro-structure, a rather unpromising substance can be transformed into a mechanically sound material. The same principle should apply in the case of cement.

In general, hydraulic cements harden by reaction with water through precipitation of near-insoluble hydration products which progressively fill the spaces between the residual cement grains, forming a cohesive matrix that binds the composite mass together. In Portland cement, the main binding agent is a colloidal calcium silicate hydrate gel of rather indefinite

Fig. 1 Tabulation of tensile strength versus work of fracture for a range of common materials.



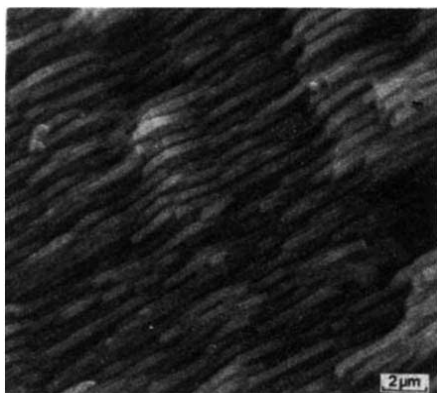


Fig. 2 Fracture surface of mother-of-pearl shell (nacre) showing intricate laminated structure of calcium carbonate (aragonite) crystals. Scanning electron micrograph. (Photograph courtesy of J.P. Northover, Oxford University.)

composition and structure. In fractured samples this gel is often seen as irregular interlocking fibres growing from the surfaces of the cement grains and interspersed with plates of crystalline calcium hydroxide, a by-product of the hydration reaction (figure 3). The mechanism by which hydration occurs and the factors responsible for the development of strength are still not fully understood (see for example ref. 4). It is recognized, however, that the main source of weakness is the residual porosity within the hydrated cement microstructure. As normally made, a hardened cement paste has a total porosity of about 20 per cent and contains pores with sizes ranging from the order of 1 mm to less than 100 Å.

The importance of the article in this issue of *Nature* (p.388) by Birchall, Howard and Kendall is the convincing demonstration that, by deliberate manipulation of the porosity, very substantial improvements can be made in the strength of cement. The flexural strength values quoted, between 60 and 70 MN m⁻², are about an order of magnitude greater than those for 'normal' cements, and even exceed previous estimates of the maximum intrinsic strength.

This has been done by application of classical theory relating to the fracture of brittle solids (Griffith's theory) and by recognition of the fact that it is porosity distribution that determines the fracture strength and not total porosity, as has hitherto been supposed. Although the total porosity of their samples remains fairly high, the increases in strength have been achieved by refinement of the porosity distribution to eliminate the large pore range — the macroscopic flaws that provide sources for easy propagation of cracks

through the microstructure. Thus, the article shows that cement is not intrinsically as weak as has been supposed and also illustrates that the improvements are attainable by fairly simple methods without the need to apply elaborate and costly treatments such as hot pressing, autoclaving and resin impregnation — methods that have been investigated by previous workers.

This may be only the beginning of the story. After physical processing to reduce the porosity to dimensions comparable with the size of the cement grains and the hydrate particles, researchers should perhaps be looking to methods of chemically modifying the hydration reaction in such a way as to promote greater cohesion within the cement paste. Previous attempts to do this have not been successful probably because of the overriding effect of macro-porosity. Chemical cross-linking by the incorporation of reactive organic molecules is one possibility; an enhancement of the natural polymerizing ability of the colloidal silicate in Portland cement is another. The chemistry is extremely complex but, nevertheless, there would seem to be considerable scope in this area. Fibre reinforcement is a well proved method of augmenting the mechanical properties of a material, particularly the fracture toughness. It has been applied with some success in the case of cement using asbestos, glass, steel and polymer fibres, but in many cases the mechanical deficiencies of the cement matrix have imposed limitations. An enhancement of the inherent strength of cement could provide considerable impetus to the devel-

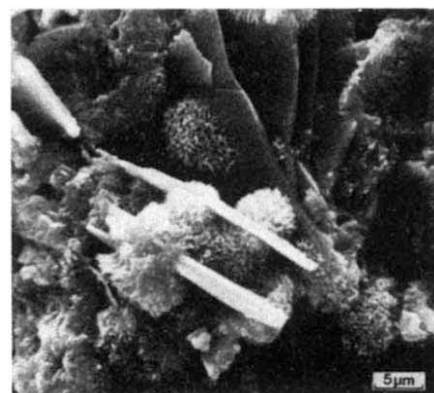


Fig. 3 Fracture surface of hardened Portland cement paste, showing fibrillar growth of calcium silicate hydrate gel from surfaces of cement grains and interspersed crystals of calcium hydroxide. Scanning electron micrograph.

opment of such composite materials.

There is still a long way to go but if the necessary improvements are possible, it is difficult to predict where cement-based materials might eventually be used. One can visualize a variety of new domestic uses and even in the area of transport. Boats have been made out of reinforced cement (ferrocement) for many years. Cement, or something rather like it, may yet fly — not as a turbine blade in a jet engine certainly, but perhaps as an important structural part of the airframe. If that day comes, then cement will have become a very respectable material indeed. □

D.D. Double is in the Department of Metallurgy and Science of Materials, University of Oxford.

Understanding interatomic forces

from Gillian Peach

THE ANALYSIS of spectra often provides the only diagnostic tool with which to study the physical conditions of a plasma or gas in the laboratory or in space. The interactions between the atoms that are emitting (or absorbing) radiation and neighbouring atoms can, however, produce a marked effect upon the observed spectrum. In 1907 R.W. Wood (*Astrophys. J.* 26, 41) wrote "I have found what appears to be indisputable evidence that the appearance, and even the apparent position, of an absorption band can be profoundly modified by the presence in the absorbing vapour of a chemically inert gas" and it is

now generally understood that the presence of a background gas can produce a broadening and shift of the individual spectral lines, and also an asymmetry in the line shape which becomes more marked as the pressure of the background gas increases. Over the last eighty years there has been an enormous number of studies of this phenomenon (see Fuhr *et al.* *NBS spec. Publ.* 366; 1972 and Suppl. 1974, 1975, 1978). In many cases, the emitter-perturber interaction is well known, and a study of line shapes is used to determine the density of perturbing particles in the emitting region; a typical example is that of hydrogen lines perturbed by protons and electrons. However, when atomic spectra are perturbed by neutral atoms or molecules, the line shapes are analysed to

1. Birchall, J.D. *Proc. Symp. Adv. in Cement Matrix Composites* (Materials Research Society, Boston, Massachusetts, 1980).
2. Ray, N.H. *Inorganic Polymers* 150 (Academic, New York, 1978).
3. Currey, J.D. *Proc. R. Soc. B* 196, 443 (1977).
4. Lea, F.M. *Chemistry of Cement and Concrete* 3rd edn, Ch.9 & 10 (Edward Arnold, London, 1970).

Gillian Peach is in the Department of Physics and Astronomy, University College London, UK.

provide information on the basic emitter-perturber interaction. This type of analysis can be extremely valuable, since it can give information on interatomic forces where one of the atoms involved is in an excited state. Studies of atom-atom scattering predominantly yield information on interactions between two atoms that are both in their ground states (see Massey *Electronic and Ionic Impact Phenomena* 2nd edn, Vol.3, Oxford University Press; 1971).

The problem of analysing these line shapes is a fascinating and difficult one. In general, the shape of the line profile is determined by the simultaneous interaction of the emitter with many perturbers; but under certain circumstances, however, the problem can be reduced to the interaction of just one perturber with the emitting atom. If the frequency separation, $\Delta\nu$, from the line centre is sufficiently large, the mean interaction energy required, $V = h\Delta\nu$, is predominantly produced by a single perturber making a close collision with the emitter. The contribution from more distant perturbers is negligible by comparison. Thus the analysis of the shape of the line wings can give detailed information on the inner region of the atom-atom potential. This approach has been used for many years (see Kuhn *Proc. R. Soc. A* **158**, 212; 230; 1937), and an excellent modern example of such an analysis is given by York *et al.* (*J. chem. Phys.* **63**, 1052; 1975). They are able to obtain the potential energy curves for the $X^2\Sigma$ and $A^2\Pi$ states of the systems Na-Ar, Na-Kr and Na-Xe. *Ab initio* calculations of the adiabatic potentials for such systems are difficult and costly, although a good deal of progress has been made in the last few years by adopting semi-empirical model or pseudo-potential methods (see Peach *J. Phys. B* **11**, 2107; 1978; Philippe *et al. J. Phys. B* **12**, 2493; 1979).

When the frequency separation from the line centre is small the mean interaction energy V is also small, and is produced by many perturbers making weak collisions with the emitter; the probability of a close collision is very low. This picture provides the basis of the impact approximation, a simple description of which is given by Peach (*Contemp. Phys.* **16**, 17; 1975). The impact approximation predicts a lorentzian profile for the line given by

$$I(\Delta\nu) = \frac{1}{\pi} \frac{w(\Delta\nu)}{(\Delta\nu - d)^2 + w^2}$$

where in this case, the half width at half maximum is $w(\Delta\nu) = w$, and is independent of frequency. The shift in the position of the line is given by d . Szudy and Baylis (*J. quant. Spectrosc. Radiat. Transfer* **15**, 641; 1975; **17**, 269, 681; 1977) have developed a scalar theory of line broadening which is less restrictive than the impact approximation. The scalar nature

of the theory means that it is assumed that the frequency shifts produced by the individual perturbers are additive, but allows for the fact that the perturbing collisions are of finite duration. Their theory predicts a first-order correction to the pure lorentzian profile in the core of the line, which is given by $w(\Delta\nu) = w + a\Delta\nu$, where a is a constant. The parameter a depends only on the average time of duration of the collision, the nature of the interactions between the emitter in both its initial and final levels and a perturber, and the temperature of the gas. Thus the line contains a dispersive component which leads to an asymmetry in the line core. Asymmetries in the far wings of lines occur for different reasons; in this case completely different regions of the emitter-perturber potentials are being sampled. Very recently the first quantitative measurements of asymmetries in line cores have been reported (Walkup *et al. Phys. Rev. Lett.* **45**, 986; 1980). They have studied the shape of the resonance lines of sodium perturbed by helium, neon, argon, krypton and xenon. They show conclusively that the lines contain a dispersive component as predicted by the theory, and obtain accurate values for the constant a . About the same time Kielkopf (*J. Phys. B* **13**, 3813; 1980) also published width, shift and asymmetry parameters for the sodium resonance lines perturbed by the same systems. His measurements are

also for similar pressures and temperature, $P = 10\text{--}600$ torr and $T = 450\text{K}$. However, his quoted errors are somewhat larger than those given by Walkup *et al.* If a pure van der Waal's potential is used for the emitter-perturber interaction, good agreement is obtained between theory and experiment for the systems Na-Ar, Na-Kr and Na-Xe. In fact this agreement is better than actually stated in the paper of Walkup *et al.* as can be shown when the full profile given by the scalar theory is used. For the systems Na-He and Na-Ne the agreement is not good; this is readily understandable since the heavier rare gases have larger dipole polarizabilities and hence the long-range part of the potential becomes relatively more important. These measurements present an exciting new challenge to the theorists. It is clear that calculations are required which use realistic interatomic potentials; on the basis of a pure van der Waal's potential, it is not possible to obtain the different values of the asymmetry parameter that are observed for the two lines. This difference gives us direct information about the difference between the $A^2\Pi_{1/2}$ and $A^2\Pi_4$ potential curves for the sodium-rare gas systems. It seems certain that more such measurements will be made, and will be used to probe regions of interatomic potentials that are not accessible in studies of the far wings of spectral lines, or in other types of atom-atom scattering experiments. □

Current issues in conservation

from Jared M. Diamond

WHEN nature reserves are looked at with the understanding gained from studies of island biogeography then the problems of avoiding major losses of the world biota seem even more difficult than at first thought. Nature reserves can be thought of as islands in a sea of man-transformed habitats. Not only will a reserve which saves only a small part of a particular habitat start out with fewer species than a larger one, it will ultimately lose a greater fraction of its species and it will lose them faster (Terborgh *Bioscience* **24**, 715; 1974). This is best illustrated by the well documented losses of species in Panama's Barro Colorado Island Reserve within this century (Willis *Ecol. Monographs* **44**, 153; 1974). Biogeographic arguments for large reserves are reinforced by genetic ones: a population with an effective size of less than 500 gradually loses the genetic variation that underlies ability to adapt, and an effective population size of less than 500 leads to inbreeding depression (Soulé and Wilcox *Conservation Biology* Sinauer; 1980). As discussion at two recent

conservation meetings* in South Africa showed, even in Kruger National Park (South Africa's largest reserve) numerous key species such as lions, ground hornbills, wild dogs, and saddle-billed storks have populations whose long-term prospect of survival is only marginal from this perspective. Biologists have long felt intuitively the importance of large reserves, but lacked the detailed arguments that have been emerging in the past decade. To many participants in the South African meetings, the detailed arguments were chilling.

The meetings revealed that biologists, park managers, and government planners have important differences in vocabulary, values, and areas of concern, and that it takes effort to bridge these communication barriers. It is clear that biologists who wish to be effective have to learn the vocabulary in which public planners express themselves, to broaden their perspective so as to be aware of social and economic constraints that concern planners, and to be specific. Part of the reason for the limited success of biologists in persuading

Jared M. Diamond is Professor of Physiology at University of California Medical School, Los Angeles.

*A symposium on the conservation of threatened natural habitats was held in Cape Town on September 11-12 and a workshop on the same topic at Tsitsikama National Park on September 14-18, 1980.

governments to establish large reserves may be the failure to offer specific arguments in terms of fraction of species expected to be lost, expected rates of losses, and likely identities of lost species in a reserve of a given size. More generally, planning and management should be part of one continuous process, with progressively more specific stages such as: formulation of a national plan of conservation priorities; selecting areas, sites, and shapes of reserves; zoning the reserve for different uses; and developing an integrated programme of research, monitoring, and management.

The last word, management, is initially anathema to many biologists. Yet a recurring theme at these meetings was that nature reserves cannot be left alone; they have to be managed. Many biologists assume, not even consciously, that a reserve need only be set aside; nature will 'manage itself' and preserve the ecosystem, and human intervention is precisely what one wants to avoid. The fatal flaw in this reasoning is: each reserve was formerly part of a much larger biogeographic unit, and it was this larger unit which 'managed itself', that is, species and habitats rarely disappeared in the whole unit; they might disappear locally but persisted or expanded elsewhere within the unit. When a reserve becomes isolated, it is doomed to change because of the loss of processes outside its boundaries that were necessary for its maintenance. The smaller the reserve, the further and faster it will change. The reasons for these inevitable changes are multiple: reserve boundaries often intersect paths of regular seasonal movements (for example, the game fence around Kruger intersecting the former west-east seasonal migrations of game); populations isolated in a reserve may be too small to be self-sustaining, though they were able to sustain themselves in the larger system; the landscape is a shifting mosaic of habitats, and a small reserve that temporarily loses a habitat essential for certain species loses those species permanently; and extinctions of a few species lead to extinctions of other species dependent on them, via effects called trophic cascades.

A spectacular example of the need for managing even large reserves is the disaster that resulted from a hands-off approach to Tsavo National Park, where reserve status led to an increase in the number of elephants producing a complete transformation of the habitat, and ultimately a crash of the elephant population. Biologists have to resign themselves: measures such as culling, re-introductions, transfers of animals, maintaining a habitat mosaic by controlled fires, and provision of water in arid areas must be standard tools of a reserve manager. The smaller the reserve, the more effort is required, but even Kruger requires constant human intervention to avoid losses of key species and transformations of habitats. We face a cruel dilemma:

without management, we will lose species and habitats; yet, only for a few species have we the biological knowledge necessary to manage well. This is not the leisurely world of academia, where one completes one's marshalling of evidence before publishing. Here, as elsewhere in the field of conservation, one must act on very limited information, often mere guesses.

Finally, the South African picture. Measured against standards of other countries, especially less wealthy countries, South Africa has done much for conservation. It has dozens of well-run reserves, including two large ones (Kruger and Kalahari Gemsbok) that are famous throughout the world. Measured against standards of what South Africa actually needs to protect its own biological heritage, there are glaring gaps (Huntley in *Biogeography and Ecology of South Africa*, 1333 (ed. M.J.A. Werger), Junk, 1978). For two habitats that covered a large fraction of South Africa, the Highveld Grassland and the Karoo, there is currently no large reserve at all. The Cape lowland

fynbos, one of the world's richest and most distinctive temperate-zone plant communities, is represented in only a few small reserves. Unless much larger areas can be protected soon, the scale of plant extinctions in the fynbos is likely to constitute one of the most serious biological losses on Earth in the next few decades. Finally, the Pafuri area of Kruger National Park is being considered for coking coal mining. On the one hand, Pafuri is ecologically the richest area of Kruger and contains many tropical species not found elsewhere in South Africa. On the other hand, the Pafuri coking coal reserves are sufficient for only about 10 or 15 years of exploitation, and South Africa already has coal reserves adequate for its own use for centuries and exports coal to other countries. The basic concept of a national park is an irrevocable assignment of land for protection of its resources in perpetuity. It would be ironic if South Africa, which played a leading role internationally in developing this concept, were now to take the lead in dismantling it at Pafuri. □

Chlorinated dioxins and the environment

from Alastair Hay

WERE it not for the bacteria and fungi to be found in the soil, the chemical industry would be having a hard time of it. Most of the organic chemicals produced by manufacturers as herbicides and pesticides are degraded by soil microorganisms. But this is not true for all chemicals. When it comes to the degradation of the polychlorinated dibenzodioxins (PCDDs) it seems that bacteria cannot cope.

The best known of the PCDDs is the tetrachlorinated isomer 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). An unwanted by-product produced in the manufacture of 2,4,5-trichlorophenols TCDD is a stable compound. A temperature in excess of 800°C is required for its thermal degradation — but, more conveniently, TCDD can be subjected to photolytic degradation provided a hydrogen donor in the form of a suitable solvent is present.

But it is not just the stability of TCDD which causes problems; the chemical is also extremely toxic with an LD₅₀ as low as 0.6 µg kg⁻¹ in the guinea pig. TCDD has also been shown to be a tetratogen and carcinogen in various animal species.

It is hardly surprising, therefore, that the discharge of 250–500g of TCDD over a populated area at the Italian town of Seveso in July 1976 should be a subject of concern. The chemical was vented directly

into the atmosphere when a 2,4,5-trichlorophenol reactor overheated.

Studies performed by the US Air Force had suggested that soil bacteria are capable of degrading TCDD and that the half life of the chemical — depending on the bacteria present — was between 230 and 320 days (Young *et al.* USAFA-TR-76-18; 1976).

At Seveso however, the half life is much longer and considerably in excess of 10 years according to results presented at a recent workshop* by A. di Domenico (Istituto Superiore di Sanita (ISS), Rome). It seems that photolytic degradation of TCDD did occur at Seveso in the weeks following the discharge of the chemical from the reactor. This breakdown took place when the chemical was on surfaces exposed to sunlight. Within a couple of months, however, the TCDD had been washed from plant surfaces and percolated into the first few centimetres of topsoil, where it has remained ever since. Di Domenico claims that the soil concentrations of TCDD in the most heavily contaminated region of Seveso have changed little in the 4 years since the accident.

Further evidence of the poor biodegrad-

*A Workshop on the Impact of Chlorinated Dioxins and related compounds on the Environment; held at the Istituto Superiore di Sanita in Rome from 22 to 24 October, 1980. The proceedings of the workshop will be published as a book by Pergamon Press.

ability of TCDD by microorganisms was provided by R. Hütter (Swiss Federal Institute of Technology, Zurich) and I. Camani (ISS, Rome). Both reported that they had not observed any microbial degradation of TCDD under laboratory conditions. A previous report by Matsumura and Benezet (*Environ. Health Perspect.* 5, 253; 1973) had suggested that there were some microbial strains which would detoxify TCDD, albeit with a low efficiency.

The fact that TCDD is so resistant to breakdown by soil microorganisms could mean that other PCDD isomers are equally difficult to degrade. Unfortunately there has been practically no investigation of the persistence of these other isomers in the environment. That they occur as environmental pollutants, however, is beyond dispute. Several laboratories in Europe have reported the presence of PCDD isomers in the fly ash from municipal and industrial incinerators. Scientists at the Dow Chemical Company claim that they can also be detected in ash from household chimneys, cigarettes and even charcoal-broiled steaks (see *Nature* 281, 619; 1979).

It is not entirely clear which compounds act as precursors of PCDDs in incinerators. According to O. Hutzinger (University of Amsterdam) the principal precursors are chlorinated phenols and polyvinyl chloride (PVC). Lignin too is a likely source, says Hutzinger, and it is conceivable that inorganic chlorine might be available in incinerators for the chlorination of dibenzodioxins. Dow scientists were the first to claim that chlorine from both inorganic and organic sources could be used in the formation of PCDDs.

Whilst there is a measure of agreement on the environmental sources of PCDD, there is a marked disagreement about the type of isomer obtained in fly ash. W. Crummett (Dow Chemicals, USA) claims that fly ash contains an equal distribution of PCDD isomers containing 4,5,6,7 and 8 chlorine atoms. The tetrachlorinated isomer 2,3,7,8-TCDD is the most toxic PCDD isomer and Dow is claiming that it is present in the ash in significant amounts. Others disagree. H.R. Buser (Swiss Federal Research Station, Wädenswil) and R.L. Harless (Environmental Protection Agency, USA) both claim that TCDD only accounts for 1% of the PCDD content of fly ash from municipal incinerators. A. Liberti (Consiglio Nazionale delle Ricerche, Rome) reported that he had not observed an equal distribution of PCDD isomers in fly ash from municipal incinerators. According to Liberti, octachlorinated dibenzodioxin — many times less toxic than TCDD — was always present in the greatest concentrations, followed closely by heptachlorinated and hexachlorinated PCDD isomers.

Can anything be done to reduce the amount of PCDs produced in large incinerators? Apparently it can. There are two alternatives; either to reduce the

amount of organic chlorinated compounds burned in incinerators, or to increase the efficiency of incineration. K. Olie (University of Amsterdam) pointed out that the PCDD content of fly ash from incinerators in Holland was 20 times higher than the content of fly ash obtained from Bratislava (Czechoslovakia) where far less PVC is incinerated. As for improvements in combustion efficiency, W. Crummett pointed out that by using extra fuel in an incinerator at the Dow Chemical Company the octachlorodibenzodioxin concentration in fly ash was reduced from 180–810 p.p.b. to 0.01–0.1 p.p.b.

High temperatures are a well recognized way of degrading PCDDs. Much less is known, however, about the biological degradation of these chemicals in mammals. TCDD is the only isomer which has been studied in any detail and Poiger and Schlatter (*Nature* 281, 706; 1979) have reported that the chemical is hydroxylated in the rat. Metabolism of TCDD also occurs in the dog and H. Poiger (University of Zurich) now believes that this metabolism is a means of detoxifying the chemical as the acute toxicity of the metabolite is at least 100 times lower than that of TCDD itself.

One of TCDD's more worrying properties is its ability to cause malformations in animals. H. Nau (Freie Universität, Berlin) reported that when TCDD was administered to pregnant mice only very low concentrations of the chemical were found in the embryo and fetus between gestational days 11 and 18. TCDD concentrations in the fetus were between 0.04 and 0.14% of the maternal dose, two orders of magnitude lower than the concentration in maternal liver — the organ in which TCDD accumulation was greatest. Repeated doses of $5 \mu\text{g kg}^{-1}$ on days 7–11 of pregnancy produced cleft palate in 65% of fetuses whereas $25 \mu\text{g kg}^{-1}$ given as a single dose on day 7 caused cleft palate in

only 16% of fetuses. There were two important conclusions to be drawn from these results said Nau. Firstly, the incidence of malformations clearly depends on the frequency of dosing, and secondly, only small amounts of TCDD are required to produce fetal abnormalities.

Detecting small amounts of TCDD in humans has always been a problem. With TCDD being lipid soluble a fat biopsy was the obvious method of choice. But biopsies have their limitations, particularly because they are uncomfortable for the donor. According to C. Rappe (University of Umea) this discomfort is no longer necessary for he has developed an assay for PCDDs using only 10 ml of blood plasma.

Workers employed in tanneries constantly handle hides which have been treated with pentachlorophenol preservatives. Rappe examined workers recently exposed to pentachlorophenol or who had been exposed 6 months previously. The results showed that pentachlorophenol is cleared from plasma when the individual is no longer exposed to the chemical. But the picture is quite different for the PCDDs which contaminate pentachlorophenol. These are far more persistent and are only cleared from plasma at a very slow rate.

Rappe's new assay will have considerable repercussions for those concerned about the health of individuals exposed to PCDDs. It will now be possible to measure the PCDD levels of workers exposed to the chemical in industry and to set more rigorous health standards. Blood measurements will be essential for interpreting clinical symptoms following individual exposure to PCDDs. In the longer term they may even assist in assessing the risk to humans of TCDD causing malformations or cancer. □

Alastair Hay is a lecturer in the Department of Chemical Pathology at the University of Leeds.

100 years ago



King of Gorontalo



Sultan of Sulu

REVIEW ARTICLE

The origin of human cancers

John Cairns*

Imperial Cancer Research Fund, Burtonhole Lane, London NW7 1AD, UK

The limited evidence available suggests that most human cancers are not caused by conventional mutagens but are more likely to be the result of genetic transpositions. Although the molecular biology of transposition is starting to be understood, the external factors that influence its frequency have not yet been studied in any detail.

DIFFERENT human populations tend to suffer from different kinds of cancer: the inhabitants of third world countries have an excess of liver cancer, those of western nations have an excess of cancer of the breast and colon, the Japanese have an excess of stomach cancer, and so on. Most of this variation must be due to varied diets, customs and environment rather than to differences in genetic constitution, because nations have been observed to undergo changes in cancer incidence from one generation to the next and migrant populations tend to take on the pattern of cancer that is characteristic of their new homes. Cancer is therefore thought to be a preventable disease.

In the western world the two commonest cancers are cancer of the skin and the lung, and each of these happens to be so strongly dependent on a single factor (that is, sunlight and smoking) that their cause could be identified without the need for any understanding of the underlying mechanism of carcinogenesis. The preventable causes of the other common cancers are less clear-cut and it may be difficult to identify them until much more is known about mechanism. Eventually the techniques of nucleic acid chemistry should allow us to itemize all the differences in nucleotide sequence and gene expression that distinguish a cancer cell from its normal counterpart, and perhaps at that point the steps involved in carcinogenesis will cease to be in doubt. But until then we have to be content with circumstantial evidence. This review discusses the evidence and makes certain deductions about the molecular biology of human cancer.

Experimental carcinogenesis

Much of cancer research is based on the reasonable premise that we should be able to learn about human cancer by studying carcinogenesis in animals. During the 65 years since coal tar was shown to be carcinogenic for the skin of rabbits, many ways have been found for producing cancer in experimental animals. The problem is to decide which of these very diverse forms of carcinogenesis is likely to be the best model for most human cancers.

The easiest and most widely studied method is to expose the chosen target in the animal to physical or chemical agents that damage DNA. Because the carcinogenic potency of such agents is fairly well correlated with their ability to cause point mutations¹⁻³, these forms of cancer are commonly assumed to arise as the result of local changes in DNA sequence (in the rest of this article, the words 'mutation' and 'mutagen' are used specifically

for such local changes). However, the process cannot be quite as simple as that. In most cases, repeated doses of the mutagen have to be administered over an appreciable fraction of the animal's lifespan, and in those few situations where a single dose of mutagen is carcinogenic the resulting cancer usually does not appear for a long time⁴. Thus the creation of a cancer cell is thought to involve a sequence of events of which perhaps only the early steps bear any direct relation to the interaction between mutagen and DNA. This is borne out by the observation that the later events can be caused by other agents ('promoters') that are not themselves mutagenic. The conventional explanation is that 'initiation' represents the production of mutants and that the subsequent steps of promotion are required because the relevant mutations either happen to be recessive (and so will only be detectable following deletion or somatic recombination) or usually involve repressor genes (and so are not expressed until the existing stock of repressor molecules has been depleted by further cell multiplication). But this type of interpretation is not without difficulties. The first step in the sequence seems to be too efficient⁵⁻⁷; in certain situations, even quite low doses of mutagen are capable of initiating the process in almost every exposed cell. Conversely, the later steps seem to be too sluggish^{8,9}; in several instances, more than 10 cell divisions may be needed following initiation to achieve expression of the new phenotype in any of the descendent cells.

Cancers can arise following infection with certain viruses. Here the explanation seems at first sight more straightforward, for it is easy to see how novel phenotypes could result from an interaction between the genes of a virus and those of the host. However, the interval between initial infection by the virus and final appearance of the cancer can be a large fraction of the animal's lifespan¹⁰, and so even here the exact biology of the carcinogenic process may sometimes be quite as obscure as in most chemical carcinogenesis.

Cancers can also be produced in animals by transplanting certain tissues, such as ovary¹¹ or embryo¹², into special sites where the cells can multiply without restraint; similarly, cancers may arise in the tissues next to implanted sheets of plastic, apparently because the cells have lost the restraining influence provided by contact with each other¹³. It seems unlikely that any conventional form of mutation can play a part in these types of carcinogenesis, particularly as the cancer cells produced by transplantation can sometimes be restored to a normal, regulated pattern of growth by transplantation back to more demanding sites¹⁴⁻¹⁶.

Lastly, a number of cancers have been observed to arise spontaneously in animals that are given unlimited food, but not

* Present address: Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115.

in animals whose intake of calories has been slightly restricted¹⁷. The mechanism of carcinogenesis by these few extra calories is completely obscure.

Those seem to be the main ways of causing cancer in animals. It is not obvious what, if anything, they have in common, and in the absence of any additional source of information we could therefore have no way of knowing which of these experimental cancers are the best models of the common human cancers; and if we do not know that, we cannot guess what class of environmental agent is likely to be the most important factor in causing human cancer.

The genetics of susceptibility to cancer

When the precise chemistry of any biological process is in doubt, it has usually been helpful to determine what kinds of mutation affect the process. This stratagem has been applied successfully in working out various metabolic pathways and in identifying many of the components involved in the synthesis of macromolecules, and it would seem to be the surest way of finding out what classes of biochemical event tend to be rate-limiting for human carcinogenesis. The method is, however, subject to certain logical restraints.

(1) When the main effect of a mutation of a multicellular animal is to alter the phenotype of one particular class of cells, the fact that it also raises the incidence of cancer in these cells may not be relevant when we come to consider the cause of other cancers. For example, some forms of inherited immunodeficiency are associated with an increased incidence of certain rare cancers arising in cells of the immune system¹⁸; obviously, this tells us little or nothing about the origin of the common cancers in other tissues.

(2) In general, mutations that increase the incidence of any cancer should be interpreted with some caution. Going back to the previous example, it is conceivable that the leukaemias and lymphomas of immunodeficient children arise by some special pathway, so that they may not even be a good model for the leukaemias and lymphomas that occur in normal subjects.

(3) The most informative mutations are usually those that stop things from happening. For example, mutations that block the induction of aryl hydrocarbon hydroxylase make mice more sensitive to the short-term toxicity and less sensitive to the long-term carcinogenicity of certain polycyclic aromatic hydrocarbons^{19,20}. If we could show that similar mutations in humans lowered the incidence of any of the common cancers, this would suggest that these compounds are important in human carcinogenesis. So far, however, such studies in human populations have been inconclusive^{21,22}.

(4) In practice, much of our information may have to come from negative evidence. For instance, if a mutation that completely blocked the metabolism of certain potential carcinogens had no effect on the incidence of cancer, this would constitute very strong evidence that these chemicals and their metabolic products were probably not important. In fact, the first case we shall consider is an example of negative evidence.

If the conventional dogma is correct and cancer is usually initiated by localized somatic mutations produced by the chemical mutagens in our environment, we would expect inherited defects in DNA repair to have a marked effect on the incidence of the common cancers. Yet this turns out not to be true. The test case is the disease called xeroderma pigmentosum (XP). This arises when both copies of any one of the genes involved in the main pathway for excision repair contain an inherited defect^{23,24}. The result seems to be equivalent to the effect of the *uvr* mutations of *Escherichia coli* and *Salmonella typhimurium*, which make these bacteria into exquisitely sensitive tester strains for most chemical carcinogens. Like such bacteria, the cells of XP patients have a defect in DNA excision that makes them up to 10 times more sensitive to the lethal²⁵⁻³² and mutagenic³³⁻³⁸ effects of most mutagens (the main exceptions being X rays and certain simple alkylating agents). As a result of

this defect, XP patients are extremely sensitive to UV light. They suffer an age-specific incidence of skin cancer that is several thousand times higher than normal and they frequently die from malignant melanoma; as these skin cancers arise on the exposed parts of the body, they are presumably the result of mutations induced by UV light in the basal cells of the skin, perhaps coupled with some radiation damage to the local immune system of the skin³⁹. Yet, even though the repair defect of XP patients extends to all the tissues that have been tested and apparently affects the response to all bulky lesions and major distortions of the double helix²⁴, it does not result in any obvious increase in the common fatal cancers (such as cancer of the lung, alimentary canal and breast). This is a striking anomaly, but it seems to have been pointed out only once before⁴⁰. For example, not one death from internal cancer was recorded in any of the published clinical reports on series of XP patients⁴¹⁻⁵⁰ summarized in Table 1; and a survey of the entire literature on XP for the past 40 years⁵¹ (which included a search through 6 years of US death certificates⁵² and therefore must comprise many hundreds of cases) has revealed only seven internal neoplasms and only two or three deaths from cancers that could not have been caused by sunlight. Now, it is possible at this point to indulge in some special pleading—by arguing, for example, that XP patients do not live long enough to experience the common internal cancers which, for some as yet unknown reason, cannot be brought forward in time as readily as skin cancers. But the most reasonable conclusion is that the commonly fatal cancers are not caused by the kind of lesions of DNA that XP patients (and bacterial *uvr* mutants) are specifically unable to repair.

That conclusion is strengthened when we consider Bloom's syndrome, an inherited disease that apparently does increase the incidence of the common cancers. Its molecular biology is not understood, but the disease is associated with a high frequency of chromosomal aberrations and exchanges^{53,54}. (Interestingly, the patients' cells do not show any very marked increase in sensitivity to various mutagens²⁴.) From a registry of cases it has been possible to calculate very crude estimates of 'life-tables' and of relative risk for carcinomas and for leukaemias and lymphomas (Fig. 1, Table 1). We see that the death rate from cancer is raised about 100-fold in each age group (Table 1). This demonstrates that it is possible to observe such increases in risk even when the number of subjects is small and their lifespan very limited; for XP, where the numbers are greater and lifespan is somewhat longer, it should have been even easier; in this sense, therefore, Bloom's syndrome serves as a control. (Two other inherited diseases, Fanconi's anaemia and ataxia telangiectasia, are associated with chromosomal instability, and they too show a high incidence of leukaemia; however, the published reports are not sufficiently detailed for the calculation of life tables.)

Comparison of the clinical records for XP and Bloom's syndrome therefore leads to the following conclusions. (1) For most people the main source of mutagenic lesions in DNA is UV light. (2) Chemical mutagens (of the kind detected in the usual tests for mutagenicity) seem unlikely to be the rate-limiting components in most human carcinogenesis. (3) Large-scale changes in the genome (such as rearrangements and deletions) seem to be more hazardous than the local changes produced by conventional mutagens.

The karyotype of human cancers

If genetic rearrangement or deletion is a critical step in the formation of most human cancers, the change might often be large enough to produce a visible alteration in chromosomal anatomy. Unfortunately, it is not easy to obtain good metaphase spreads for the cancer cells in solid tumours, and so most studies of karyotype have been confined to the leukaemias⁵⁵ and to a few solid tumours, such as meningiomas⁵⁶ and gliomas⁵⁷, which contain neoplastic cells that grow readily *in vitro*. In these

diseases, visible chromosomal rearrangements have proved to be more the rule than the exception, and for many cancers the karyotype has actually been observed to become progressively abnormal with time. It is therefore clear that at least the later stages in development of cancers are usually associated with large-scale changes in the genome.

For obvious reasons, much less is known about the earlier stages of carcinogenesis. About 90% of chronic myeloid leukaemias show a translocation from chromosome 22 to 9, which creates the characteristically shortened 22 known as the Philadelphia chromosome⁵⁸; in several instances⁵⁹⁻⁶¹ the translocation has been observed to be present quite early in the sequence of events leading to the leukaemia, suggesting that it was primarily this alteration that allowed the precancerous stem cell to multiply without restraint and gradually colonize most of the bone marrow. (The exceptional cases of chronic myeloid leukaemia that do not have the translocation run a distinctly less favourable clinical course⁶², and so they can probably be classified as a different kind of cancer, brought on by a different sequence of events.) Similarly, patients with ataxia telangiectasia are often found to be harbouring an expanding clone of lymphocytes with a translocation affecting chromosome 14 (refs 63, 64) and the lymphatic leukaemia which commonly occurs in these patients apparently develops within these clones. (Interestingly, myelomas and one class of lymphomas may also show translocations to chromosome 14 (ref. 65).) Those are the clearest examples where a particular change in karyotype appears to be an essential, early step in carcinogenesis. Of course, it is possible that still earlier in each of those sequences some crucial step occurred that made inevitable the subsequent emergence of cells bearing these particular rearrangements. For example, benign meningiomas have often been observed to be diploid in their early stages⁶⁶, and it is only later that they tend to be dominated by a characteristic cell that has lost one chromosome 22 (ref. 56); and there are certain kinds of cancer, such as acute myeloid leukaemia, that sometimes maintain an apparently normal karyotype throughout the whole course of the disease⁶⁶.

These studies of karyotype can therefore be interpreted in one of two ways. The obvious conclusion is that large chromosomal rearrangements are usually the crucial steps in carcinogenesis, but that there are exceptions and in certain classes of cell the critical changes either are not rearrangements or are on too small a scale to be detected by chromosomal banding techniques. It is, however, possible to argue that the cellular environment within any growing cancer is very abnormal and applies strong selection pressure for very unusual phenotypes, and these are most easily generated by large-scale rearrangements of the genome; the observed changes in karyotype could therefore be trivial secondary events that occur after all the rate-limiting steps of carcinogenesis have been completed.

The evolutionary analogy is worth pursuing a little further. The normal architecture and programmes of cell renewal in multicellular animals seem to be so arranged that the various multiplying stem cells do not, as a rule, compete with each other for territory⁶⁷. For example, it is clear from studies of X-chromosome mosaicism that adult human tissues are made up of a large number of very small families of cells which therefore cannot be undergoing clonal selection^{68,69}; similarly, the haematopoietic stem cells that circulate in the blood of mice are capable of establishing themselves in the marrow and forming clones of descendants, but they will only do so if the resident stem cells occupying the available sites in the marrow have been killed⁷⁰. So one of the crucial steps in the development of a cancer must be the change that allows a stem cell to compete favourably with its neighbours and form an expanding clone within which further natural selection can then take place. We should therefore consider the possibility that the formation of a cancer proceeds by the same mechanism as the evolution of a new species. In other words, we might expect normal non-cancerous cells to be like species that are not evolving quickly: they will accumulate local alterations in base sequence and

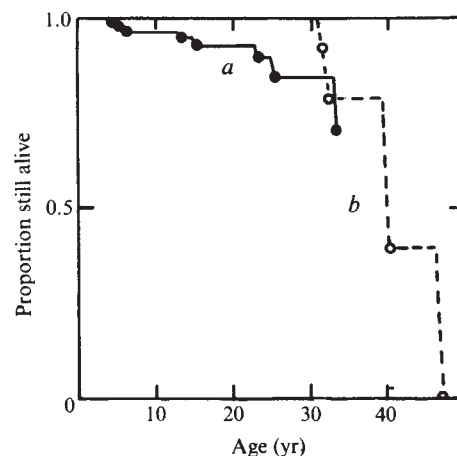


Fig. 1 Life tables for patients with Bloom's syndrome. The calculations have been based on a collection of 89 cases (J. German, unpublished), 12 of whom have died of cancer. *a*, Proportion who have not yet died of leukaemia or lymphoma; *b*, proportion who have not yet died of carcinoma.

amino acid sequence due to forced or unforced errors in DNA replication, but these will tend to be neutral and have little effect on phenotype⁷¹. In contrast, a family of cancer cells should be like a species that is undergoing a rapid evolution in morphology. Such evolution is now thought to be due mainly to pleiotropic changes in the regulation of gene expression that are the result of large-scale alterations in gene arrangement and chromosome constitution^{72,73}. In short, the analogy with evolution would predict that cancer should usually be due to gene rearrangements rather than to the local changes in sequence produced by conventional mutagens. Before considering what is known about the molecular biology of such rearrangements, it is worth examining the other half of the proposition, namely that carcinogenesis by conventional mutagenesis is not making a major contribution to the national incidence of cancer.

Mutagenesis and cancer

It seems clearly established that localized lesions in DNA can be carcinogenic: for example, pyrimidine dimers must be the cause of the skin cancers produced when animals are irradiated with UV light, because the cancers can be prevented by photo-reversal with visible light (a reaction that seems to be specific for dimers)⁷⁴; similarly, the formation of *O*⁶-alkylguanine in DNA is apparently the basis for carcinogenesis by certain alkylating agents because the cancers tend to be confined to the tissues that are unable to repair that particular lesion^{75,76}. So the issue is not simply whether conventional mutagens are capable of causing cancer, but whether they reach our cells in sufficient quantity to make a major contribution to our present national incidence of cancer. In short, it is a question of the balance between dose of mutagens and the efficiency of our various pathways for DNA repair.

Although skin cancer due to UV light is the commonest cancer in western nations, it is seldom fatal and accounts for only 1% of all cancer deaths. But when there is a defect in the relevant DNA repair pathway (as in XP) the age-specific incidence of skin cancer can be raised several thousandfold and life-expectancy drop to about 20; this shows the extent to which an efficient system of repair is protecting us against one of the major hazards of our environment. Similarly, the repair pathway that handles the lesions produced by X-ray irradiation seems to be equally effective, because only about 1% of all cancer deaths are believed to be attributable to the natural sources of ionizing

radiation⁷⁷. Presumably there has always been strong selection pressure for adequate methods of DNA repair, and so it is hardly surprising to find such low levels of risk from agents like UV light and X rays, which have not changed in intensity for millions of years.

There are reasons for thinking that our exposure to the chemical mutagens in our environment is now actually somewhat less than it used to be. The argument runs as follows⁷⁸. Experimental animals tend to show a rather limited range of responses when mutagens are added to their diet; a few, very reactive chemicals produce cancer of the oesophagus and stomach, but most mutagens produce cancer of the liver. In humans, however, liver cancer accounts for only 1% of the cancer deaths in countries like the United States and Europe; oesophageal cancer is not very common outside certain parts of Africa, Iran and China; and stomach cancer is on the decline in all industrialized nations. So it seems likely that when affluence brings us the opportunity to choose what we eat, we tend to avoid foods heavily contaminated with mutagens and are therefore able to stay well within the capabilities of our DNA repair pathways. Actually, even in undeveloped countries the level of exposure to chemical carcinogens seems to be within most people's capacity for repair; in Africa, for example, the liver cancer caused by aflatoxin is largely confined to people chronically infected with hepatitis B virus⁷⁹, which indicates that although aflatoxin is known to be one of the most potent carcinogens for certain experimental animals it is not present in a high enough concentration even in Africa to produce many cancers in normal people.

So much for the 'natural' mutagens in our environment. Fortunately, the novel unnatural mutagens that we have met in the first half of this century as a result of our industrialization seem to be no more powerful than the natural ones⁸⁰. Occupational cancers probably account for less than 5% of all cancer deaths⁸¹; and the number attributable to mutagenesis is actually less than this because a large proportion of today's occupational cancers are due to asbestos, which is negligibly mutagenic⁸² although it is known to produce chromosomal abnormalities⁸³. These numbers are, of course, trivial compared with the effect of tobacco. This has proved to be the outstanding carcinogen of the

twentieth century and is now responsible for about a third of all cancer deaths. But it is not at all clear that these cancers can be classified as due to mutagenesis. Although cigarette smoke condensate contains mutagens⁸⁴, it is much more powerful as a promoter than an initiator in experimental carcinogenesis^{85,86}. Furthermore, if tobacco were mainly acting as a mutagen, there should have been some reports of lung cancer in XP patients; after all, strong sunlight (which in normal people is about as carcinogenic as heavy smoking) can produce skin cancer in these patients by the time they are 7 years old, and so smoking should have given some of them lung cancer by the time they are 30.

All these arguments are indirect, but they do suggest rather strongly that local changes in DNA sequence, produced by conventional mutagens, are probably making only a minor contribution to our national death rates.

Genetic transposition as a cause of cancer

In the last decade, the routes leading to genetic diversity have become much better understood^{87,88}. Diffusion of heritable information is now known to occur not only by 'legitimate' recombination between homologous regions of different genomes, but also by 'illegitimate' recombination between largely nonhomologous regions (which can be in different parts of the genome or even in different individuals). The second of these processes is equivalent to a chromosomal rearrangement and allows the shuffling of certain modules of information ('transposons') so that new combinations of genes can be tested for survival advantage. The process depends on the presence of certain short sequences in the DNA that are the substrates for various recombinational enzymes ('transposases'), and it represents a much more drastic form of genetic variation than localized changes in base sequence because it can alter the expression of whole regions of the genome.

Transpositions can be site specific (for example, the changes in surface antigens of *Salmonella* and certain trypanosomes) or nonspecific (such as the spread of antibiotic resistance between different bacteria); they are usually solitary, but they can occur in clusters; and their frequency can vary from 1 per cell division down to 10^{-10} . A transposon can be the substrate for a transposase encoded elsewhere in the genome or it can make its own specific transposase. As the result of moving to a new site, a transposon can start or stop being transcribed; but the main effect can be not on itself so much as on neighbouring genes, which can be turned on or off or made unstable by its presence.

Transposition is therefore a very special form of genetic regulation because it can be programmed to occur at any predetermined rate, can be concerned with any desired number of alternative states, and can be made virtually irreversible. For example, the switch in the mating type of yeasts is the result of a transposition which occurs at a rate that is itself under separate genetic control⁸⁹; but the change has all the characteristics of a step in differentiation because it concerns the function of a large number of genes but normally is confined to one particular branch of the cell lineage (the mother cells, rather than the buds)⁹⁰. It is not yet clear whether transposition is important in vertebrate development. As the result of certain experiments on nuclear transplantation, there has been a tendency in recent years to regard the nuclei of the somatic cells of vertebrates as totipotent and therefore not fundamentally different from the germ line⁹¹. However, the fact remains that although it is easy to produce a normal animal using the nucleus taken from a fertilized egg or an early blastula cell, no one has ever succeeded in producing a normal animal using the nucleus taken from a somatic cell of an adult. It is therefore quite conceivable that changes in gene arrangement do occur during vertebrate development, especially as it is now known that antibody diversity is generated by specific transpositions that occur during the differentiation of B lymphocytes^{92,93}.

As pointed out in the earlier sections of this review, there are good reasons for believing that the steps leading to human cancer are more likely to be transpositions than localized

Table 1 Deaths from internal cancers

Age	Patient years	Bloom's syndrome* obs./exp.†		Patient years	Xeroderma pigmentosum‡ obs./exp. (excl. skin)
		Leukaemias and lymphomas	Carcinomas		
0-4	340	1/0.013	0/0.017	640	0/0.059
5-14	640	3/0.020	0/0.019	886	0/0.055
15-24	327	2/0.011	0/0.016	450	0/0.036
25-34	108	2/0.005	2/0.013	213	0/0.035
35-44	16		1/0.008	109	0/0.038
45-54	3		1/0.005	54	0/0.107
Totals		12/0.127			0/0.330

The observed values were obtained from various clinical reports. For Bloom's syndrome the estimates are probably fairly reliable, because the condition is very rare and so the fate of each case is usually followed rather closely; of these 89 cases, 17 are known to have died (12 from cancer). Because XP is much less rare, cases tend to be reported but not followed up, and so the existing literature may be used to give some idea of their incidence of cancer but will give an underestimate of their death rate. During the limited period when these 140 cases were observed, 71 had had skin cancer (which was fatal in 8 cases), 1 had had a brain tumour and 1 had chronic myeloid leukaemia.

*89 cases (J. German, unpublished).

†Expected values were calculated from the current age-specific cancer death rates in England and Wales.

‡140 cases reported in refs 41-50.

changes in sequence. It is therefore important to find out what classes of external agents or features of cellular behaviour raise the frequency of such transpositions. Interestingly, conventional mutagens often appear to have no effect. Indeed, it was the inability of any mutagen to raise the spontaneous reversion rate of certain forward mutations in *E. coli* that initially marked out these mutations as being unusual⁹⁴ and eventually led to the discovery that they were due to the movement of insertion sequences; similarly, the movement of certain transposable elements in yeast has been found to occur 'spontaneously' but does not appear to be strongly catalysed by any of the common mutagens⁹⁵. However, rates of transposition can be influenced by external factors. For example, transpositions are only a minor cause of mutation in rapidly multiplying bacteria⁹⁶, but they are a major source of forward mutations in stationary cultures⁹⁷; the growth of *Drosophila* cells *in vitro* leads to the proliferation of certain repetitive sequences⁹⁸; the frequency of the transpositions that produce variegation in *Drosophila* and in *Antirrhinum* petals is inversely related to temperature^{99,100}; and the integration of a tumour virus can be associated with transpositions in the neighbouring host DNA¹⁰¹. Lastly, many mutagens are known to cause sister chromatid exchanges and other chromosomal interactions in mammalian cells, although it

is not always clear to what extent these gross changes are functionally equivalent to genetic transpositions.

The assay of mutagens using bacterial tester strains has been one of the most fruitful practical results of molecular genetics. Its success stems from the fact that the chemistry of DNA is the same for all forms of life and so the chemistry of mutagenesis (and of repair) also tends to be the same. The molecular biology of transposition promises to be more idiosyncratic. The transposable genetic elements were discovered by McClintock¹⁰² in 1950 and their possible role in carcinogenesis was first discussed by Temin¹⁰³ in 1970, but it is still too early for there to have been much systematic molecular biological investigation of the factors that trigger transpositions, either in prokaryotes or in eukaryotes. The rate of movement of most transposons will have been subject to evolutionary pressures and many transposons will therefore have acquired their own individual controls. Thus it may not be a simple matter to devise a general assay for the factors that drive the carcinogenic transpositions.

I thank Errol Friedberg, Richard Peto, Jeffrey Miller, Julian Gross and Bryn Bridges for useful advice and discussion, James German and Kenneth Kraemer for unpublished information about patients with Bloom's syndrome and XP, and David Roberts for a literature search for clinical reports on XP.

- Meselson, M. & Russell, K. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 1473-1481 (Cold Spring Harbor Laboratory, New York, 1977).
- Ames, B. N. & Hooper, K. *Nature* **274**, 19-20 (1978).
- Bartsch, H. et al. in *Molecular and Cellular Aspects of Carcinogen Screening Tests* (eds Montesano, R., Bartsch, H. & Tomatis, L.) 179-241 (IARC, Lyon, 1980).
- Laerum, O. D. & Rajewsky, M. F. *J. natn. Cancer Inst.* **55**, 1177-1187 (1975).
- Mondal, S. & Heidelberger, C. *Proc. natn. Acad. Sci. U.S.A.* **65**, 219-225 (1970).
- Parodi, S. & Brambilla, G. *Mutat. Res.* **47**, 53-74 (1977).
- Kennedy, A. R., Fox, M., Murphy, G. & Little, J. B. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Kakunaga, T. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1334-1338 (1978).
- Barrett, J. C. & Ts'o, P. O. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3761-3765 (1978).
- Allison, A. C., Chesterman, F. C. & Baron, S. J. *natn. Cancer Inst.* **38**, 567-572 (1967).
- Furth, J. J. *natn. Cancer Inst.* **8**, 7-16 (1947).
- Stevens, L. C. *Dev. Biol.* **21**, 364-382 (1970).
- Karp, R. D. et al. *J. natn. Cancer Inst.* **51**, 1275-1285 (1973).
- Brinster, R. L. *J. exp. Med.* **140**, 1049-1056 (1974).
- Mintz, B. & Illmensee, K. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3585-3589 (1975).
- Papadogiannou, V. E., McBurney, M. W., Gardner, R. L. & Evans, M. J. *Nature* **258**, 70-73 (1975).
- Tannenbaum, A. & Silverstone, H. *Adv. Cancer Res.* **1**, 451-501 (1953).
- Good, R. A. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1026-1032 (1972).
- Kouri, R. E., Rattie, H. & Whitmore, C. E. *J. natn. Cancer Inst.* **51**, 197-200 (1973).
- Nebert, D. W. *J. natn. Cancer Inst.* **64**, 1279-1290 (1980).
- Atlas, S. A., Vesell, E. S. & Nebert, D. W. *Cancer Res.* **36**, 4619-4630 (1976).
- Paigen, B. et al. *Cancer Res.* **37**, 1829-1837 (1977).
- Cleaver, J. E. & Bootsma, D. A. *Rev. Genet.* **9**, 19-38 (1975).
- Friedberg, E. C., Ehmann, U. K. & Williams, J. I. *Adv. Radiat. Biol.* **8**, 85-174 (1979).
- Kleijer, W. J., Lohman, P. H. M., Mulder, M. P. & Bootsma, D. *Mutat. Res.* **9**, 517-523 (1970).
- Cleaver, J. E. *Mutat. Res.* **12**, 453-462 (1971).
- Slor, H. *Mutat. Res.* **19**, 231-235 (1973).
- Maher, V. M., Birch, N., Otto, J. R. & McCormick, J. J. *J. natn. Cancer Inst.* **54**, 1287-1294 (1975).
- Maher, V. M., Ouellette, L. M., Mittlestat, M. & McCormick, J. J. *Nature* **258**, 760-763 (1975).
- Fornace, A. J. & Kohn, K. W. *Biochim. biophys. Acta* **435**, 95-103 (1976).
- Day, R. S., Scudiero, D. & Dimattina, M. *Mutat. Res.* **50**, 383-394 (1978).
- Andrews, A. D., Barrett, S. F. & Robbins, J. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1984-1988 (1978).
- Maher, V. M., Ouellette, L. M., Curren, R. D. & McCormick, J. J. *Nature* **261**, 593-595 (1976).
- Maher, V. M., McCormick, J. J., Grover, P. L. & Sims, P. *Mutat. Res.* **43**, 117-138 (1977).
- Arlett, C. F. & Harcourt, S. A. in *DNA Repair Mechanisms* (eds Hanawalt, P. C., Friedberg, E. C. & Fox, C. F.) 633-636 (Academic, New York, 1978).
- Maher, V. M., Dorney, D. J., Mendrala, A. L., Konze-Thomas, B. & McCormick, J. J. *Mutat. Res.* **62**, 311-323 (1979).
- Myhr, B. C., Turnbull, D. & DiPaolo, J. A. *Mutat. Res.* **62**, 341-353 (1979).
- Glover, T. W., Chang, C. C., Trosko, J. E. & Li, S. S. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3982-3986 (1979).
- Bridges, B. & Strauss, G. *Nature* **283**, 523-524 (1980).
- German, J. in *Proc. 6th int. Congr. Radiation Research* (eds Okada, S. et al.) 496-505 (Tokyo University Press, 1979).
- Reese, A. B. & Wilber, I. E. *Am. J. Ophthalmol.* **26**, 901-911 (1943).
- Berlin, C. & Tager, A. *Dermatologica* **116**, 27-35 (1958).
- El-Hefnawi, H., El-Nabawi, M. & Rasheed, A. *Br. J. Derm.* **74**, 201-213 (1962).
- El-Hefnawi, H. & Rasheed, A. *J. Egypt. med. Ass.* **46**, 985-1006 (1963); **49**, 419-434 (1966).
- Lemaire, M. & Gaumond, E. *Can. med. Ass. J.* **92**, 406-412 (1965).
- Siegelman, M. H. & Sutow, W. W. *J. Pediatr.* **67**, 625-630 (1965).
- Sambasivan, T. S., Siddappa, K. & Indrakumar, S. V. *Indian J. Derm.* **11**, 79-81 (1966).
- Robbins, J. H., Kraemer, K. H., Lutzner, M. A., Festoff, B. W. & Coon, H. G. *Ann. intern. Med.* **80**, 221-248 (1974).
- Takebe, H. et al. *Cancer Res.* **37**, 490-495 (1977).
- Pawsey, S. A., Magnus, I. A., Ramsay, C. A., Benson, P. F. & Giannelli, F. Q. *J. Med. new Ser.* **48**, 179-210 (1979).
- Kraemer, K. H. in *Clinical Dermatology* (eds Demis, D. J., Dobson, R. L. & McGuire, J.) (Harper & Row, New York, in the press).
- Miller, R. W. *Pediat. Res.* **3**, 389-397 (1969).
- German, J., Archibald, R. & Bloom, D. *Science* **148**, 506-507 (1965).
- German, J. in *Chromosomes and Cancer* (ed. German, J.) 601-617 (Wiley, New York, 1974).
- Mittelman, F. & Levan, G. *Cancer Genet. Cytogenet.* **1**, 29-32 (1979).
- Zankl, H., Zang, K. D. *Cancer Genet. Cytogenet.* **1**, 351-356 (1980).
- Mark, J., Westermark, B., Pontén, J. & Hugosson, R. *Hereditas* **87**, 243-260 (1977).
- Rowley, J. D. *Nature* **243**, 290-293 (1973).
- Canellos, G. P. & Whang-Peng, J. *Lancet* **ii**, 1227-1229 (1972).
- Baccarani, M., Zaccaria, A. & Tura, S. *Lancet* **ii**, 1094 (1973).
- Verhest, A. & van Schoubroeck, F. *Lancet* **ii**, 1386 (1973).
- Whang-Peng, J., Canellos, G. P., Carbone, P. P. & Tijo, J. H. *Blood* **32**, 755-766 (1968).
- Cohen, M. M., Shaham, M., Dagan, J., Shmueli, E. & Kohn, G. *Cytogenet. Cell Genet.* **15**, 338-356 (1975).
- McCaw, B. K., Hecht, F., Harnden, D. G. & Teplitz, R. L. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2071-2075 (1975).
- Mark, J., Dahlenfors, R. & Ekedahl, C. *Cancer Genet. Cytogenet.* **1**, 39-56 (1979).
- Mark, J. *Adv. Cancer Res.* **24**, 165-222 (1977).
- Cairns, J. *Nature* **225**, 197-200 (1975).
- Gartler, S. M., Gandini, E., Hutchison, H. T., Campbell, B. & Zechhi, G. *Ann. hum. Genet.* **35**, 1-7 (1971).
- Feder, N. *Nature* **263**, 67-69 (1976).
- Mickleth, H. S., Ogden, D. A., Evans, E. P., Ford, C. E. & Gray, J. G. *Cell Tissue Kinet.* **8**, 233-248 (1975).
- Wilson, A. C., Carlson, S. S. & White, T. J. A. *Rev. Biochem.* **46**, 573-639 (1977).
- Wilson, A. C., Bush, G. L., Case, S. M. & King, M. C. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5061-5065 (1975).
- Bush, G. L., Case, S. M., Wilson, A. C. & Patton, J. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3942-3946 (1977).
- Hart, R. W., Setlow, R. B. & Woodhead, A. D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5574-5578 (1977).
- Goth, R. & Rajewsky, M. F. *Proc. natn. Acad. Sci. U.S.A.* **71**, 639-643 (1974).
- Margison, G. P. & Kleihues, P. *Biochem. J.* **148**, 521-525 (1975).
- Jablson, S. & Bailar, J. C. *Preventative Med.* **9**, 219-226 (1980).
- Peto, R. in *Origins of Human Cancer* (eds Hiatt, H. H., Watson, J. D. & Winsten, J. A.) 1403-1428 (Cold Spring Harbor Laboratory, New York, 1977).
- Larouze, B. et al. *Lancet* **ii**, 534-538 (1976).
- Doll, R. *Proc. R. Soc. B205*, 47-61 (1979).
- Higginson, J. & Muir, C. S. J. *natn. Cancer Inst.* **63**, 1291-1298 (1979).
- Huang, S. L. *Mutat. Res.* **68**, 265-274 (1979).
- Sincock, A. & Seabright, M. *Nature* **257**, 56-58 (1975).
- Kier, L. D., Yamasaki, E. & Ames, B. N. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4159-4163 (1974).
- Roe, F. J. C., Salaman, M. H. & Cohen, J. *Br. J. Cancer* **13**, 623-633 (1959).
- van Duuren, B. L., Sivak, A., Segal, A., Orris, L. & Langseth, L. *J. natn. Cancer Inst.* **37**, 519-526 (1966).
- Nevers, P. & Saedler, H. *Nature* **268**, 109-115 (1977).
- Calos, M. P. & Miller, J. H. *Cell* **20**, 579-595 (1980).
- Hicks, J. B., Strathern, J. N. & Herskowitz, I. in *DNA Insertion Elements, Plasmids and Episomes* (eds Bukhari, A. I., Shapiro, J. A. & Adhya, S. L.) 457-462 (Cold Spring Harbor Laboratory, New York, 1977).
- Strathern, J. N. & Herskowitz, I. *Cell* **17**, 371-381 (1979).
- Gurdon, J. B., Laskey, R. A. & Reeves, O. R. *J. Embryol. exp. Morph.* **34**, 93-112 (1975).
- Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Saedler, H. & Starlinger, P. *Molec. gen. Genet.* **100**, 178-189 (1967).
- Roeder, G. S., Farabaugh, P. J., Chaleff, D. T. & Fink, G. R. *Science* **209**, 1375-1380 (1980).
- Calos, M. P., Johnsrud, L. & Miller, J. H. *Cell* **13**, 411-418 (1978).
- Arber, W. et al. *Cold Spring Harb. Symp. quant. Biol.* **43**, 1197-1208 (1978).
- Potter, S. S., Brorine, W. J., Dunsin, P. & Rubin, G. M. *Cell* **17**, 415-427 (1979).
- Gowen, J. W. & Gay, E. H. *Science* **77**, 312 (1933).
- Harrison, B. J. & Fincham, J. R. S. *Hereditas* **19**, 237-258 (1964).
- Botchan, M., Stringer, J., Mitchison, T. & Sambrook, J. *Cell* **20**, 143-152 (1980).
- McClintock, B. *Proc. natn. Acad. Sci. U.S.A.* **36**, 344-355 (1950).
- Temin, H. M. *Perspectives Biol. Med.* **14**, 11-26 (1970).

ARTICLES

Thermodynamic stability and kinetics of perovskite dissolution

H. Wayne Nesbitt, G. Michael Bancroft, William S. Fyfe,
Suresh N. Karkhanis & Akio Nishijima

Centre for Chemical Physics and Department of Geology, University of Western Ontario, London, Canada N6A 3K7

Shigemitsu Shin

National Chemical Laboratory for Industry, Tsukuba, Japan

Perovskite, a SYNROC host mineral for nuclear wastes, is thermodynamically unstable in natural waters and in association with common minerals. Leach experiments demonstrate that CaTiO_3 (perovskite), SrTiO_3 and BaTiO_3 are as reactive as some silicate glasses below 100°C , but leach much more slowly than glasses above 100°C .

THE use of titanates and zirconates as hosts for nuclear waste in the SYNROC process has recently been proposed^{1,2}. Three minerals are used as hosts—perovskite (CaTiO_3), Ba-hollandite ($\text{BaAl}_2\text{Ti}_6\text{O}_{16}$) and zirconolite ($\text{CaZrTi}_2\text{O}_7$). SYNROC relies heavily on geological and geochemical observations in selecting stable host minerals. Although the SYNROC minerals are not thermodynamically compatible with some siliceous rocks, the minerals are considered likely to be thermodynamically stable in the presence of water. Moreover, it has been reported that these minerals are kinetically stable in extreme hydrothermal conditions (up to 900°C and 5,000 bar), and are thus vastly more stable to leaching than borosilicate glasses^{1,2}.

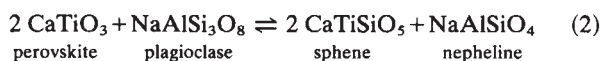
We report here thermodynamic calculations and leaching kinetics from 25°C to 300°C on the simplest of the three SYNROC minerals, perovskite, and its Sr and Ba analogues. Perovskite is suitable for such a thermodynamic and kinetic study because its dissolution reactions are simple, and the thermodynamic data for perovskite and its reaction products are known for a wide range of temperatures. By comparison with glasses, the leach rate of perovskite and its analogues is very insensitive to temperature, due to an interesting combination of the thermodynamics and kinetics of MTiO_3 ($\text{M} = \text{Ca}, \text{Sr}, \text{Ba}$) dissolution.

Thermodynamic stability

We first consider the thermodynamic stability of perovskite in the presence of the most common silicates, and then consider the three most likely reactions of perovskite in natural waters.

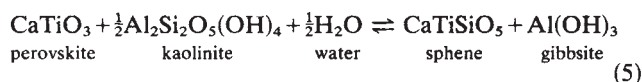
The plagioclase and alkali feldspars are the most abundant minerals in the Earth's crust followed by quartz (and polymorphs) and orthopyroxenes. These three mineral groups contain 50 mol % SiO_2 or more and constitute between 80 and 90% of all minerals present in the crust³. Perovskite stability in the presence of these siliceous minerals must be considered for they may react directly or indirectly (via a fluid) with the titanate.

Consider the reactions:



The thermochemical data for these phases⁴ have been used to calculate the free energies for each of the reactions (ΔG_R^0) at various temperatures. The free energies of all reaction are negative at 25°C indicating that sphene is stable relative to perovskite in the presence of quartz, feldspars and orthopyroxenes. The ΔG_R^0 values become progressively more positive at higher temperatures but all values remain more negative than -10 kJ even at 900°C . The incompatibility of perovskite and quartz, orthopyroxene and particularly feldspar severely restricts the repository or vault type if this potential for perovskite alteration is to be minimized. Virtually all plutonic and most volcanic rocks would have to be discarded for they generally contain feldspars and/or orthopyroxene. The only igneous rocks suitable would be low Si, high Ca rocks such as those given by Carmichael *et al.*⁵. Alternatively, non-siliceous rocks such as evaporite deposits and carbonates would minimize this source of potential alteration.

Kaolinite [$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$] is one of the most common clay minerals produced by weathering of aluminous silicates (such as feldspars) near the Earth's surface⁶; hence it is an important constituent of many muds, mudstones, shales and argillaceous rocks. Perovskite stability in its presence can be determined from:



The free energy of the reaction is negative at low temperature indicating that clay minerals can supply sufficient SiO_2 to destabilize perovskite. The result demonstrates that even shales and argillaceous rocks would be inadequate vaults in that they provide the potential for perovskite alteration. Ringwood *et al.*^{1,2} recognized the incompatibility of quartz and perovskite and they emphasized that thermodynamically compatible geological materials should be used for vaults. The above calculations emphasize that the choice of suitable host rocks is much more restricted than they suggested—if thermodynamic stability dictates the choice.

Experimental evidence shows that perovskite does react with common minerals at measurable rates. Schuiling and Vink⁷ studied the hydrothermal stability of perovskite in the presence of siliceous minerals. The results agree with the above calculations. Furthermore, they converted at least 5% of the perovskite to sphene in only 3 weeks (700°C , 1,000 bar pressure), as this amount is required for positive detection by X-ray diffraction. Their studies demonstrate that perovskite is unstable in the presence of quartz, feldspars and orthopyroxene.

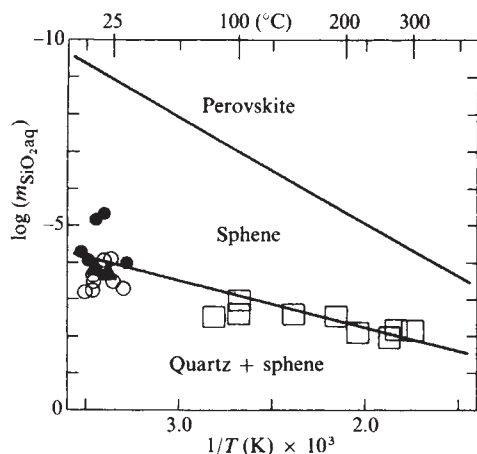
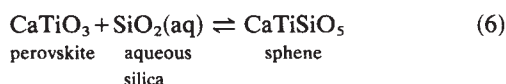


Fig. 1 $\log m_{\text{SiO}_2(\text{aq})}$ plotted against $1/T$ for reaction (6). The m_{SiO_2} , T values are plotted for groundwaters emanating from dunites and peridotites ($\bullet$), serpentinites ($\circ$), limestones (Δ) and hydrothermal waters ($\square$). These all plot very close to the lower solid line representing solutions saturated with quartz.

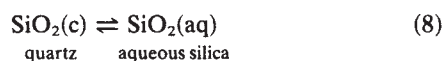
Aqueous solutions are almost ubiquitous in the crust, even to great depths⁸ and of all interactions involving perovskite, reaction with subsurface waters are most likely; consequently the stability of perovskite in natural groundwaters must be investigated. Two reactions limit the stability of perovskite. One is



where

$$\log K_6 = -\log[\text{SiO}_2(\text{aq})] \quad (7)$$

Square brackets denote activities. The $\Delta G_{R(6)}^0$ and $\log K_6$ values were calculated and the equilibrium aqueous silica concentrations plotted against the inverse of temperature (Fig. 1). Groundwaters emanating from dunites, peridotites^{9,10}, serpentinites and limestones as well as hydrothermal waters have been plotted and all fall within the sphene stability field. Perovskite is unstable in these waters. Quartz solubility is given by



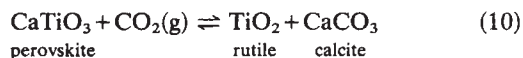
where

$$\log K_8 = \log[\text{SiO}_2(\text{aq})] \quad (9)$$

The lower solid line (Fig. 1) represents solutions saturated with respect to quartz.

Carbon dioxide is dissolved in most natural groundwater¹¹ and is one of the most abundant gases dissolved in hydrothermal

waters⁹. Carbon dioxide in natural waters (carbonic acid) may enhance the thermodynamic and kinetic instability of perovskite as shown below

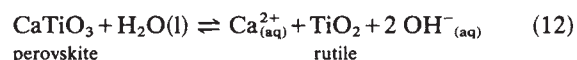


For this equilibrium relation

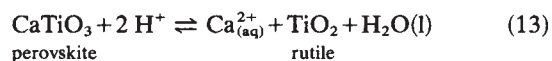
$$\Delta G_{R(10)}^0 = -RT \ln K_{(10)} = +RT \ln P_{\text{CO}_2} \quad (11)$$

Both the $\Delta G_{R(10)}^0$ and the equilibrium carbon dioxide pressure (P_{CO_2}) have been evaluated. The logarithm of the equilibrium partial pressure has been plotted against the reciprocal of temperature (Fig. 2). Some near-surface groundwaters from limestones¹¹ and from some ultramafic rocks¹² plot well within the rutile + calcite field. White *et al.*¹¹ compiled typical groundwaters emanating from a large variety of igneous, metamorphic and sedimentary rocks and all have P_{CO_2} values sufficiently high to alter perovskite to the much more stable (and common) assemblage, calcite + rutile. The calculations demonstrate that perovskites are generally unstable in CO_2 -bearing groundwater and hydrothermal waters. Limestones and dolomites would, therefore, be unsatisfactory repository materials at least in environments where the temperatures are $< 200^\circ\text{C}$.

A third reaction also influences perovskite stability:



and when recast using hydrogen ions is



Both the $\Delta G_{R(13)}^0$ and the equilibrium constant (K_{13}) have been calculated where

$$\log K_{13} = \log [\text{Ca}^{2+}]/[\text{H}^+]^2$$

Square brackets denote activities; activities of the solids and $\text{H}_2\text{O}(\text{l})$ are taken as unity and the thermodynamic data is from Robie *et al.*⁴ and Helgeson¹⁰.

Equilibrium diagrams which include the relations of equations (6) and (13) have been constructed for 25°C . Figure 3 delineates the stability of the three Ti-rich solids (plus amorphous silica) in the presence of aqueous solutions. Perovskite is stable only in groundwaters of very high $[\text{Ca}^{2+}]/[\text{H}^+]^2$ values ($> 10^{18.3}$) and very low aqueous silica values ($< 10^{-9}$). Solution compositions typical of near-surface ground waters in igneous rocks (ref. 12, Tables 1, 2 and 3) have been plotted (Fig. 3). Waters derived from limestones, dolostones and evaporites (data from ref. 12, Tables 6, 7 and 8) plot in the same region. As Fig. 3 shows, natural waters are too rich in $\text{SiO}_2(\text{aq})$ and have

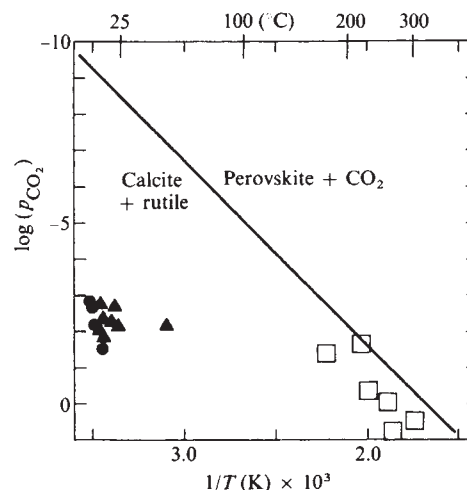


Fig. 2 $\log P_{\text{CO}_2}$ plotted against $1/T$ for reaction (10). P_{CO_2} , T values are plotted for near-surface groundwaters from limestones (Δ), ultramafic rocks ($\bullet$) and hydrothermal solutions ($\square$).

Table 1 Weight loss and Ca^{2+} leached into solution* for perovskite single crystal

Conditions	Weight loss (μg)	Total Ca^{2+} (μg)	Leach rate [†] ($\text{kg m}^{-2} \text{s}^{-1}$)
I 300 $^\circ\text{C}$, 1 week	500	430	8.9×10^{-9}
II 300 $^\circ\text{C}$, 2 weeks	100	54	5.6×10^{-10}
III 100 $^\circ\text{C}$, 1 week	220	94	2.2×10^{-9}

* This single crystal from Magnet Cove, Arkansas, weighed $\sim 120 \text{ mg}$ and had a $\sim 0.8 \text{ cm}^2$ geometric surface area. The electron probe analysis of this sample showed it to be homogeneous and gave the analyses: 38.9% CaO , 42.3% TiO_2 , 7.8% Nb_2O_5 , 5.9% FeO , 1.2% Al_2O_3 and 0.8% SiO_2 (total 96.9%). The high Fe and Nb analyses are common for perovskites in carbonatites from Magnet Cove¹³. The remainder of the analysis is most likely rare earths¹³. A few extremely small inclusions of apatite were also seen, and analysed by the probe.
[†] Based on Ca^{2+} leached into solution. $\text{kg m}^{-2} \text{sec}^{-1} = 1.16 \times 10^{-4} \text{ g cm}^{-2} \text{ day}^{-1}$.

$[Ca^{2+}]/[H^+]^2$ values too low to be equilibrated with perovskite. Furthermore these variables are not greatly influenced by the nature of the rock from which they were collected. This is not surprising because the composition of a groundwater is the net result of reactions with all minerals and rocks it has contacted, not just the last materials encountered. For example, groundwaters which have flowed through an igneous rock may have picked up appreciable $SiO_2(aq)$ by dissolution of feldspar. The $SiO_2(aq)$ will not necessarily be precipitated if the groundwater contacts limestones or evaporites. We conclude that there is probably no 'inert' vault material for perovskite because the most likely reaction is between perovskite and groundwaters, the groundwaters having contacted and reacted with various rock types.

Experimental

Four different types of samples were studied: a perovskite single crystal from Magnet Cove, Arkansas^{13,14}, perovskite grains from Wolgedee Hill, Western Australia, and synthetic $SrTiO_3$ and $BaTiO_3$. All samples showed excellent X-ray powder diffraction patterns characteristic of the titanate (see footnotes to the tables for other details).

Leaching experiments were performed using stainless steel autoclaves and distilled water. Because we expected very low leaching rates at low temperatures, we began our leaching experiments at 300 °C. The Magnet Cove sample was polished with 0.5 μm alumina on a micropolishing cloth before leaching for one week at 300 °C in 60 ml water, repolished, and then leached for two weeks in identical conditions. After repolishing, the sample was leached for one week at 100 °C.

The other three samples were placed in 10 ml of water in smaller stainless steel autoclaves—usually for one week. Several leach runs were made on the Australian perovskite at 100 °C, before using the same samples for leaching at 300 °C, and then 100 °C again. The samples were then ground to expose fresh surface before leaching for several weeks at 25 °C. The Ba and Sr titanates were leached at 25 °C, 100 °C, 200 °C and 300 °C often for two or three weeks (one week at a time). They were not polished or scraped between runs.

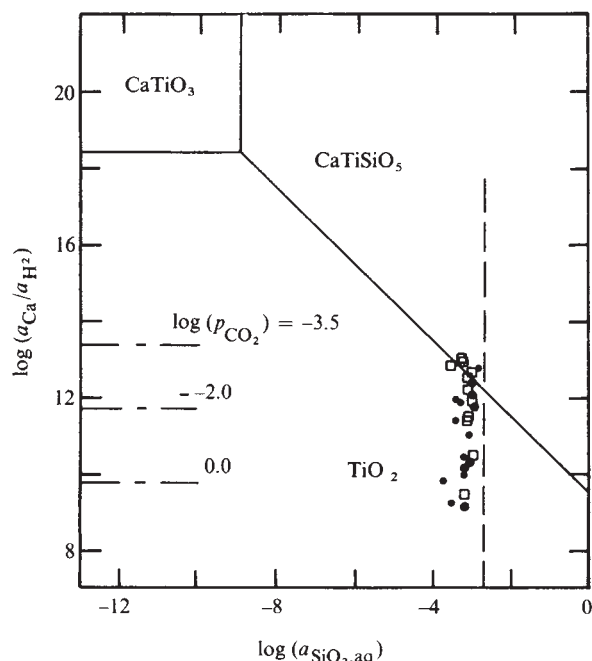


Fig. 3 $\log [a_{Ca^{2+}}/a_{H^+}^2]$ plotted against $\log a_{SiO_2(aq)}$ combining equations (6) and (13) at 25 °C. Solution compositions are plotted for basalts and gabbros ($\square$), rhyolites and granites ($\bullet$). The equilibrium line delineating the stability fields of sphene and rutile is obtained by subtracting the $\log k$ values of reaction (6) and (13). The dash-dot lines represent calcite solubility at the stipulated partial pressures of CO_2 .

After leaching, the samples were examined by a Philips 501 SEM and an ESCA 36 spectrometer, and the solution was analysed for Ca, Sr, Ba and sometimes Ti by atomic absorption spectroscopy.

Results and discussion

In the first set of experiments with the Magnet Cove single crystal, we were able to obtain both an accurate weight loss and total amount of leached Ca^{2+} (Table 1). Although we did not deliberately exclude CO_2 , the weight loss and Ca^{2+} in solution are consistent qualitatively with reaction (13) taking place. According to this equation the total weight of leached Ca^{2+} should be $\sim 72\%$ of the weight loss. Note that the leach rate in experiment II ($5.6 \times 10^{-10} \text{ kg m}^{-2} \text{ s}^{-1}$ or $4.8 \times 10^{-6} \text{ g cm}^{-2} \text{ day}^{-1}$) is much lower than that in I ($8.9 \times 10^{-9} \text{ kg m}^{-2} \text{ s}^{-1}$) while the leach rate at 100 °C if anything is higher than the latter 300 °C leach rate. Scanning electron micrographs on the leached crystal indicated new crystals on the surface, but these were not large enough to identify.

Four separate $\sim 0.5\text{-g}$ samples of the Western Australia perovskite grains were then leached for seven days in 10 ml of distilled water at 100 °C. The average amount of Ca^{2+} ion leached into solution was $78 \pm 20 \mu\text{g}$. An approximate surface area of $10 \pm 2 \text{ cm}^2$ was obtained by microscopic examination, and the leach rate becomes $1 \times 10^{-10} \text{ kg m}^{-2} \text{ s}^{-1}$, substantially lower than that obtained on the single crystal. Two of the perovskite grain samples were then leached at 300 °C and then 100 °C for one week, yielding leach rates of $1 \times 10^{-11} \text{ kg m}^{-2} \text{ s}^{-1}$ and $6 \times 10^{-11} \text{ kg m}^{-2} \text{ s}^{-1}$ respectively. After crushing the samples to expose fresh surface, the samples were leached at 25 °C for twenty consecutive weeks, yielding $10 \pm 2 \mu\text{g}$ Ca per week, and a leach rate of $< 1 \times 10^{-11} \text{ kg m}^{-2} \text{ s}^{-1}$.

The reason for the difference in leach rates between the crystal and grains is not clear. It is not due to the tiny apatite inclusions (Table 1) in the perovskite crystal because our leach tests on apatite show similar leach rates to perovskite. We could find no evidence for extensive weathering or other impurities, such as calcite, in the crystal. We cannot, of course, absolutely rule out very minor, more soluble impurities affecting the leach rates. However, since glass leach rates decrease greatly with increasing SiO_2 , it does not seem unreasonable that the low TiO_2 perovskite crystal (42.3% TiO_2) has a substantially higher leach rate than the high TiO_2 perovskite grains (58.01% TiO_2).

Table 2 summarises the leach data for the $SrTiO_3$ and $BaTiO_3$ synthetic samples. Several interesting points emerge. First, as for the above perovskite samples, the leach rate decreases markedly with leaching time. For $SrTiO_3$ at 100 °C, the leach rate (in $\text{kg m}^{-2} \text{ s}^{-1}$) decreases from 1.5×10^{-8} for a one day leach to 1.9×10^{-9} for a two week leach (Table 2). Second, as for the two perovskite samples, the leach rates for both the Sr and Ba samples at all temperatures from 25 to 300 °C are within an order of magnitude of each other. Third, as for the perovskite grains, the leach rate at 100 °C is significantly higher than the leach rates at the lower and higher temperatures. Fourth, the leach rates on the synthetic perovskite analogues are similar to the crystalline $CaTiO_3$ results (Table 3), but considerably larger than those for the Western Australia perovskite grains.

The leach rates for perovskite and the Sr and Ba analogues are summarized in Table 3. Except for our Western Australia perovskite sample, the results were obtained on disk-like materials in static tests in similar volumes of water. Our results are at least qualitatively consistent with SYNROC leaching results¹⁵ on powder samples¹⁶. The Ca leach rates reported earlier¹⁶ (10 days leaching) are within an order of magnitude for all temperatures between 25 °C and 450 °C, and the leach rates drop markedly with time¹⁶. For example, as for our Sr sample at 100 °C (Table 2), the Ca leach rate from SYNROC at 100 °C drops by an order of magnitude from a 1 day to a 10 day leach. The leach rates reported¹⁶ (8×10^{-12} to $6 \times 10^{-11} \text{ kg m}^{-2} \text{ day}^{-1}$ from 25 °C to 450 °C) are certainly lower than our crystal and disk leach rates by as much as two orders of magnitude, but are rather close to our leach rates on perovskite grains. Powder

Table 2 Sr and Ba leach rates in perovskite analogues

Sample	Leaching T(°C)	Leaching time (days)	Surface area (cm ²)	1st week		2nd week		3rd week	
				Sr or Ba (μg)	Leach rate (kg m ⁻² s ⁻¹)	Sr or Ba (μg)	Leach rate (kg m ⁻² s ⁻¹)	Sr or Ba (μg)	Leach rate* (kg m ⁻² s ⁻¹)
SrTiO ₃	25	7	2.1	300	2.3×10^{-9}	190	1.5×10^{-9}		
	100	1	2.1	265	1.5×10^{-8}				
	100	7	2.2	500	3.7×10^{-9}	170	1.3×10^{-9}	120	9.3×10^{-10}
	100	14	0.9	205	1.9×10^{-9}				
	200	7	2.4	100	7.0×10^{-10}	45	3.1×10^{-10}	ND†	$< 6 \times 10^{-11}$
	300	7	2.0	60	5.0×10^{-10}	35	2.9×10^{-10}	ND	$< 6 \times 10^{-11}$
BaTiO ₃	25	7	2.1	140	1.0×10^{-9}				
	100	7	2.1	510	4.1×10^{-9}	450	4.1×10^{-9}	350	2.7×10^{-9}
	200	7	2.0	110	9.2×10^{-10}	110	9.2×10^{-10}	90	7.4×10^{-10}
	300	7	1.8	ND	$\leq 4 \times 10^{-10}$	ND	$\leq 4 \times 10^{-10}$	ND	$\leq 4 \times 10^{-10}$

Samples were prepared by mixing the alkaline earth carbonate with TiO₂ for 3 h, calcining at 900 °C for 3 h, remixing for 3 h, calcining at 1,000 °C for 3 h, remixing, pressing and molding; before recalcining at 1,150 °C for 24 h. These samples weighed about 0.5 g, had geometric surface areas of ~2 cm², and had densities at least 90% of the theoretical values.

* Leach rates based on amounts of Sr or Ba in solution. Reproducibility is ±40%.

† ND, concentrations of Ba or Sr were below atomic absorption detection limits.

leach rates (using BET surface areas) can be as much as two orders of magnitude lower than disk leach rates using geometric surface areas^{17,18}.

Table 3 compares the leach rates for the titanates with those of rhyolite¹⁹, and borosilicates²⁰. Typical of a process controlled by kinetic factors, the glass leach rates increase dramatically with temperature. As a result, the leach rates for glasses at 200 °C or 300 °C are orders of magnitude larger than the titanates. However, at 100 °C and below the leach rates are similar; and at 25 °C it seems that, except for the Western Australia sample, the titanate leach rates are in fact larger than those for rhyolite¹⁹ or higher melting (1,300 °C) borosilicates (K. B. Harvey, personal communication).

The unusual temperature dependence of the titanate leach rate is due to an interesting combination of the thermodynamics and kinetics. At 100 °C and 300 °C, the equilibrium Ca²⁺ concentrations for equation (13) at neutral pH are ~170 M (~7 × 10⁶ μg per g or p.p.m.) and ~5 × 10⁻³ M (~200 p.p.m.) respectively. The relatively large leach rates at low temperatures are at least partially due to the large driving force provided by the extremely large equilibrium concentrations (equation (5.20)

in ref. 8). The solubility of Ca(OH)₂, drops dramatically with increasing temperature from over 1,000 p.p.m. Ca²⁺ at 20 °C, to 370 p.p.m. at 100 °C to 20 p.p.m. (or 200 μg in 10 ml) at 250 °C (ref. 21). The solubility of the Sr and Ba hydroxides are likely to be similar. Indeed, SrTiO₃ and BaTiO₃ leached at 300 °C become coated with small crystals which we have observed by scanning electron microscopy. Preliminary X-ray diffraction, and electron spectroscopy for chemical analysis (ESCA) results provide evidence that these crystals are indeed the alkaline earth hydroxides which have reprecipitated on the surface.

Thermodynamics are clearly very important in controlling the leach rate. Moreover, it is not valid to do experiments at high temperatures and extrapolate to low temperature to obtain mineral leach rates using the kinetic model of glass leaching¹.

Perovskite geochemistry

Titanium is an accessory element of many common minerals²² including sheet silicates (biotite), amphiboles (hornblende) and pyroxene (augite). It is an essential constituent of only a few minerals, such as sphene (CaTiSiO₅), ilmenite (FeTiO₃), rutile and its polymorphs (TiO₂) and perovskite (CaTiO₃). Ilmenite, sphene and rutile are not susceptible to alteration in sedimentary and hydrothermal environments and these resistate minerals may accumulate in sediments to be mined as mineral sands²³. Despite the apparent stability of Ti-rich phases, one mineral, perovskite, may not be a resistate. Geological information supports this suggestion. Perovskite is a rare mineral formed only in rocks of unusual composition. In contrast, ilmenite, sphene, and rutile are common minerals found in rocks of extremely diverse composition, suggesting that they are stable relative to perovskite in most geological environments. Rocks containing perovskite are of plutonic and volcanic origin (ref. 22, p. 423), displaying exceptionally low SiO₂ and high CaO concentrations⁵. They have been generated in very high temperature–pressure environments, probably being produced in the mantle rather than in the Earth's crust (ref. 5, p. 526). We may therefore expect the rocks and constituent minerals to be unstable in upper crustal environments. Furthermore these rocks contain, in addition to perovskite, leucite (KAlSi₂O₆), kalsilite (KAlSiO₄) and nepheline (NaAlSi₃O₈) all of which are less stable than feldspars, in the presence of natural waters¹⁰. Their survival at the Earth's surface may result more from the volume of rock to be altered and from the arid climates in which they are found (Wyoming; Western Australia; southern Spain) than from their resistance to weathering and groundwater solutions. Perovskite is also observed in limestones thermally altered by granitic intrusions²² but again the environment is unusual and very restricted. Minerals formed in such a setting may well be very unstable in other, more common, environments. From the geological evidence we conclude that perovskite will be stable only in unique geological

Table 3 Leach rates of glasses and titanates (kg m⁻² s⁻¹)

	25 °C	100 °C	200 °C	300 °C
CaTiO ₃ (crystal)	—	2.2×10^{-9}	—	8.9×10^{-9}
CaTiO ₃ (grains)*	$< 1 \times 10^{-11} \dagger$	1×10^{-10}	—	$1 \times 10^{-11} \dagger$
SrTiO ₃	2.3×10^{-9}	3.7×10^{-9}	7.0×10^{-10}	5.0×10^{-10}
BaTiO ₃	1.1×10^{-9}	4.0×10^{-9}	9.2×10^{-10}	3.5×10^{-10}
Rhyolite‡	1.0×10^{-10}	4.1×10^{-9}	3.7×10^{-8}	1.5×10^{-7}
Na borosili- cates ¶	$\geq 6 \times 10^{-12}$ $10^{-9} - 10^{-10}$	$\geq 1.2 \times 10^{-9}$ 1.4×10^{-8}	$3.5 \times 10^{-10} §$	

All results are given for the leaching of disks in the first week unless otherwise noted. The titanate leach rates are determined from Ca²⁺, Sr²⁺ or Ba²⁺ in solution, while the glass results are determined from weight losses. The leach rates drop off drastically with time. For example, rhyolite and borosilicate leach rates can be in the 10⁻¹² range after 3 month's leaching at 25 °C.

* These Western Australia grains had diameters of ~0.6 mm. The probe analysis gave: 37.66% CaO, 58.01% TiO₂ (total 95.67%) with small amounts of Na, Sr, Nb, Fe and other trace elements (see ref. 14).

† These leach rates will be comparatively low, because the sample had been leached before.

‡ Ref. 19.

§ Leach rates from Na (in solution not corrected for congruent dissolution).

¶ 30 day static leach (K. B. Harvey, personal communication). These glasses melted at ~1300 °C.

¶ Ref. 20.

settings. Our thermodynamic and kinetic studies strongly support this suggestion.

Conclusions

Contrary to earlier suggestions, we have found that perovskite is thermodynamically unstable in most rocks, and in natural waters. Perovskite appears to be kinetically far more stable than glasses at high temperatures ($T \gg 100^\circ\text{C}$). However, at the more important temperatures of $\leq 100^\circ\text{C}$, the comparison depends on the type of both glass and perovskite under consideration. For example, our best perovskite leach rate is lower than low melting borosilicates at all temperatures; but is comparable to, or greater than, the higher melting ($1,300^\circ\text{C}$) borosilicate at 25°C . Our synthetic samples leach more rapidly at 25°C than even some low melting borosilicates. Obviously, perovskite is not 'vastly' more stable kinetically than glass and silicates

generally as has been claimed^{1,2}, and much more work is required to understand fully the kinetic stability and decomposition pathways of perovskite and other SYNROC minerals.

Ideally, minerals used as nuclear waste hosts should be thermodynamically stable. A mineral which is thermodynamically unstable and has a low laboratory leach rate, could well leach much faster in the field due to chemical or biological catalysts. Considerable efforts should now be made to examine the thermodynamics of mineral and glass solubilities considering the backfill and repository or vault materials.

We thank the Atomic Energy of Canada Ltd for financial support and Dr H. H. Schloessin and Dr B. Mason for perovskite samples. This paper was presented in part at the Workshop on Alternate Nuclear Waste Forms and Interaction in Geologic Media, held at Gatlinburg in May 1980.

Received 7 August; accepted 23 December 1980.

1. Ringwood, A. E., Kesson, S. E., Ware, N. G., Hibberson, W. & Major, A. *Geochim. J.* **13**, 141–165 (1979).
2. Ringwood, A. E. & Kesson, S. E., *Proc. int. Symp. Ceramics in Nuclear Waste Management*, CONF-790420 174–178. (US Department of Energy, Technical Information Center, Oakridge, 1979).
3. Mason, B. *Principles of Geochemistry*, 100–104 (Wiley, New York, 1966).
4. Robie, R. A., Hemingway, B. S. & Fisher, J. R. *Bull. U.S. geol. Surv.* **1452**, (1978).
5. Carmichael, I. S. E., Turner, F. J. & Verhoogen, J. *Igneous Petrology* (McGraw-Hill, New York, 1974).
6. Krauskopf, K. B. *Introduction to Geochemistry* Ch. 4 (McGraw-Hill, New York, 1967).
7. Schuiling, R. D. & Vink, B. W. *Geochim. cosmochim. Acta* **31**, 2399–2411 (1967).
8. Fyfe, W. S., Price, N. J. & Thompson, A. B. *Fluids in the Earth's Crust* (Elsevier, Amsterdam, 1978).
9. Ellis, A. J. in *Geochemistry of Hydrothermal Ore Deposits* Ch. 11 (ed. Barnes, H. L.) (Holt, Rinehart and Winston, New York, 1967).
10. Helgeson, H. C. *Am. J. Sci.* **267**, 729–804 (1969).
11. White, D. E., Hem, J. D. & Waring, G. A. *Data of Geochemistry* 6th edn, Ch. F (Geological Survey Professional Paper 440-F, 1963).

12. Nesbitt, H. W. & Bricker, O. P. *Geochim. Cosmochim. Acta* **43**, 403–409 (1978).
13. Erickson, R. L. & Balde, L. V. *Geol. Surv. Profess. Pap.* No. 425 (1963).
14. Mason, B. N. *Z. Dept Sci. Ind. Res. Bull.* **218**, 114–120 (1977).
15. Reeve, K. D. et al. *Proc. Waste Management '80 Symp.* (University of Arizona, 1980).
16. Tewhey, J. D. *Workshop on Alternate Waste Forms and Interactions in Geologic Media*, Gatlinburg (1980).
17. Hench, L. L. & Clark, D. E. *J. Non-Crystal. Solids* **28**, 83–105 (1978).
18. Ross, W. A., Turcotte, R. P., Mendel, J. E. & Rusin, J. M. *Proc. int. Symp. Ceramics in Nuclear Waste Management* CONF-790420, 52–56 (US Department of Energy, Technical Information Center, Oakridge 1979).
19. Karkhanis, S. N., Bancroft, G. M., Fyfe, W. S. & Brown, J. D. *Nature* **284**, 435–437 (1980).
20. Boulton, K. A., Dalton, J. T., Hall, A. R., Hough, A. & Marples, J. A. C. *Proc. int. Symp. Ceramics in Nuclear Waste Management* CONF-790420, 248–255 (US Department of Energy, Technical Information Center, Oakridge, 1979).
21. Linke, W. F. *Solubilities of Inorganic and Metal Organic Compounds* (American Chemical Society, 1965).
22. Deer, W. A., Howie, R. A., & Zussman, J. *An Introduction to the Rock Forming Minerals* (Longmans, London, 1966).
23. Hutton, C. O. *Bull. geol. Soc. Am.* **61**, 635 (1950).

The Norse merman as an optical phenomenon

W. H. Lehn

Department of Electrical Engineering, University of Manitoba, Winnipeg, Canada R3T 2N2

I. Schroeder

60 Thatcher Drive, Winnipeg, Canada R3T 2L3

Mediaeval Norse writings that describe the merman are considered accurate observations of a natural phenomenon. Images of common sea mammals, severely distorted by strong, non-uniform atmospheric refraction, fit the mediaeval descriptions remarkably well. Three examples are presented: computer-generated images of a killer whale and a walrus, and a photograph of a suitably distended boulder. The mediaeval correlation of a merman sighting with the advent of storms is also verified.

MEDIAEVAL manuscripts contain occasional passages, apparently deriving from invention or superstition, for which a basis can be found in observation or experience. However, some texts, such as those that describe the merman, have defied positive identification. In the seventeenth century it was first suggested that the merman was simply a dugong or manatee¹, a view that was generally accepted by the end of the nineteenth century². Although this interpretation is very unsatisfactory, it has remained the most prevalent theory³.

In general, modern scientists have avoided the study of such anomaly sightings, partially because it is regarded as a non-soluble problem, and partially because it can be so easily distorted⁴.

This article considers the merman, as conceived by twelfth and thirteenth century authors. The natural source of the merman legend will be identified, and the high level of sophistication attained by the Norse in observing and correlating natural phenomena will be demonstrated. Finally, the concept of the merman will be traced through its subsequent evolution.

Mediaeval accounts

The *King's Mirror*, written in the mid-thirteenth century, is noted for its highly accurate descriptions of sea mammals and

other natural phenomena⁵. The section dealing with the North Atlantic contains only three items that assume an aspect of the marvellous: the *hafgerdingar* and the merman, found in the Greenland Sea, and the *hafgufa*, or kraken, in the Iceland Sea⁶. The explanation of the *hafgerdingar* (sea fences) as a visual effect created by anomalous atmospheric refraction⁷ suggests that the merman be reconsidered.

The author of the *King's Mirror* describes the merman as:

"This monster is tall and of great size and rises straight out of the water. . . . It has shoulders like a man's but no hands. Its body apparently grows narrower from the shoulders down, so that the lower down it has been observed, the more slender it has seemed to be. But no one has ever seen how the lower end is shaped. . . . No one has ever observed it closely enough to determine whether its body has scales like a fish or skin like a man. Whenever the monster has shown itself, men have always been sure that a storm would follow."

He proceeds to discuss its mate, the mermaid, which differs from the merman in that it possesses heavy hair, breasts, large webbed hands, and a fish's tail, but also "rarely appears except before violent storms".

An earlier account of the merman is found in the *Historia Norvegiae*, written⁸ about 1170:

"there is the whale, and there is the merman, the largest wild beast, but without head and tail, thus it is just like a tree-trunk when it leaps up and down, and it does not show itself without its foretelling danger for seafarers."

These two passages, the only descriptions of the merman that have survived from mediaeval times, will be analysed here. First the merman's shape will be considered as an optically distorted image; then the correlation between such a sighting and the advent of storms will be briefly assessed.

Atmospheric optics

The atmosphere occasionally produces optical distortion strong enough to make ordinary objects unrecognizable, even at fairly close distances. Such distortions have been suggested as the source for many of the sightings, both ancient and modern, of lake and sea monsters⁹. Common sea mammals, so distorted, fit the mediaeval merman descriptions.

The underlying optical theory, here briefly summarized, is given elsewhere¹⁰⁻¹². Light rays travelling nearly horizontally follow paths whose vertical curvature κ depends on the distribution of atmospheric density ρ with elevation z . The curvature is given by

$$\kappa = \frac{\beta}{1 + \epsilon\rho} \frac{d\rho}{dz}$$

where $1 + \epsilon\rho$ is the refractive index of air ($\epsilon = 0.000226$). The density is related to the more easily measured absolute temperature T by¹³

$$\rho = \frac{\beta p_0}{T(z)} \exp \left[-g\beta \int_0^z \frac{dz'}{T(z')} \right]$$

where p_0 is the surface pressure, g the acceleration due to gravity, and $\beta = 3.49 \times 10^{-3}$, a constant of proportionality. Temperature inversions, with their associated stable atmosphere, cause a downward refraction of light rays. Irregularities

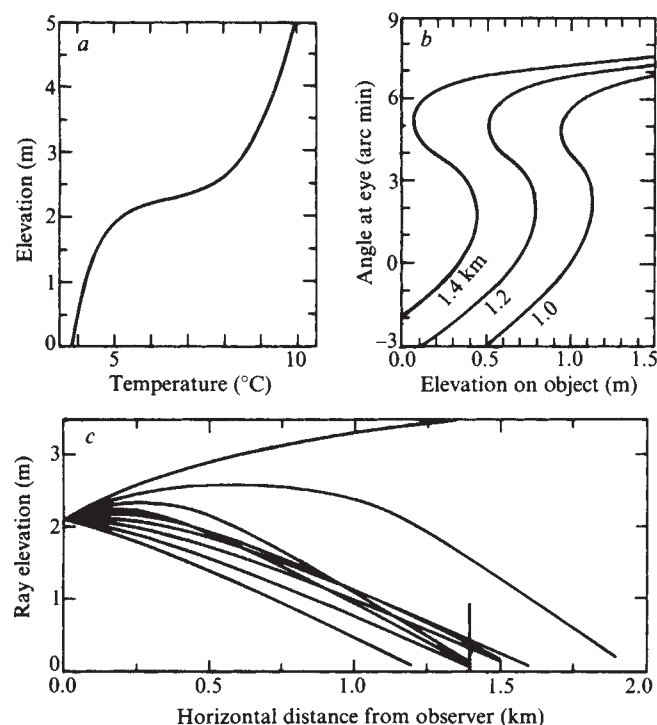


Fig. 1 *a*, Temperature profile used in Example 1. *b*, Transfer characteristics generated by the profile of *a*, for object distances of 1.0, 1.2 and 1.4 km. Eye elevation is 2.1 m. *c*, Paths described by light rays in temperature conditions of *a*. The uppermost ray enters the eye at an angle of +9°; the lowest with -3°. The angular increment between successive rays is 1.5°; the spacing of 0.5° used to plot *b* is too dense to produce a meaningful plot at this scale. The ordinate measures ray elevation above the (curved) surface of the sea. The location and height of the whale's head (see Fig. 2*a*) is indicated by the vertical bar at 1.4 km.

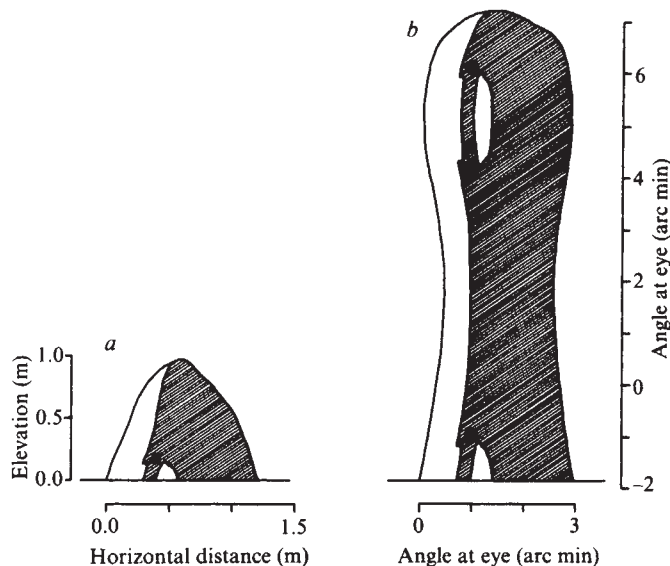


Fig. 2 *a*, A killer whale, spy-hopping. *b*, The distorted image of the whale, seen from a distance of 1.4 km in the temperature conditions of Fig. 1*a*. If the same temperature profile were to prevail for some distance beyond the whale, then the *hafgerdingar* would be present; the apparent horizon would be above the top of the image, at an elevation of +10°, about 10 km from the observer. For a normal atmosphere these values are -2.5° and 5.7 km.

in the temperature profile, especially thermoclines, create the irregularities in refraction necessary for severe image distortion.

Computer programs developed for ray tracing^{13,14} have been used to identify temperature profiles that create merman images. The reconstructions that follow are based on inversions with thermoclines situated a few metres above the sea.

Example 1: A killer whale projecting its head vertically out of the water is an excellent source for merman images. Such activity, called 'spy-hopping', is common among cetaceans of many sorts^{15,16}. It is also specifically mentioned by Shackleton¹⁷ as a hunting procedure used by killer whales¹⁸.

In the image-construction process, several quantities can be manipulated: shape of temperature profile, elevation of observer's eye and distance from observer to object. After some computer experimentation with all of these elements, the following typical conditions were chosen to illustrate the effect. Figure 1*a* shows the temperature profile: a moderate inversion of total strength 7.5°C and with a thermocline at elevation 2.2 m. The atmosphere is assumed laterally homogeneous over the short ranges involved, so that this temperature profile prevails everywhere between object and observer. If the observer's eye is now placed at an elevation of 2.1 m, the light rays entering the eye follow the paths shown in Fig. 1*c*. The elevations at which these rays intersect an object plane at some distance from the eye are plotted as a transfer characteristic in Fig. 1*b*. These curves determine the nature of the image perceived by the observer. The eye assumes each incoming ray to be straight, and projects it back on to the object plane to give an (apparent) image point. A complete set of curved rays, as in Fig. 1*b*, thus causes a redistribution of all the object points on the object plane. As the ray bending is in a purely vertical plane no lateral image distortion is present.

The original and distorted images are compared in Fig. 2*a* and *b*, respectively, for an object distance of 1.4 km. Figure 2*a* is adapted from a photograph of a killer whale¹⁹. The distorted image in Fig. 2*b* is unlikely to be recognized as a whale.

Space restrictions mean that the many possible variations of the image cannot be displayed. With changes in the determining parameters, the shape can appear more or less slender; the midsection can become quite a narrow 'neck' or 'waist'; and the white eye-spot may not be imaged in the 'head'.

Example 2: A similar image-constructing process was carried out for a walrus. As this is a smaller animal, a shorter range was

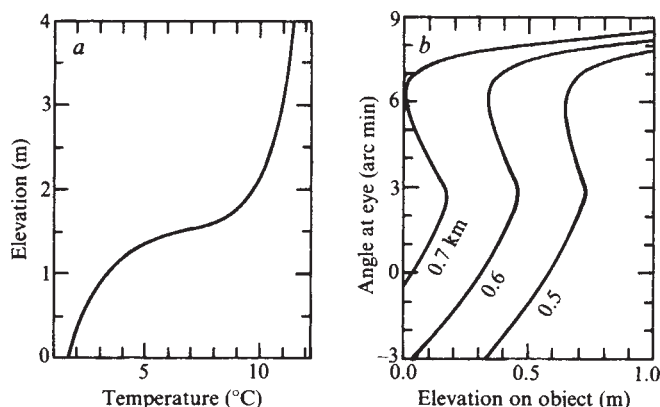


Fig. 3 *a*, Temperature profile used in Example 2. *b*, Corresponding transfer characteristics, for object distances of 0.5, 0.6 and 0.7 km. Eye elevation is 1.45 m.

chosen (0.7 km) to have a sufficient horizontal angle subtended at the eye. A steeper temperature profile was required to produce the necessary refraction over the reduced range. Figure 3a shows the temperature profile, an inversion of strength 11 °C with a thermocline at about 1.5 m elevation. The transfer characteristics corresponding to an eye elevation of 1.45 m are shown in Fig. 3b.

Figure 4 compares the original with the distorted image. A walrus photographed in its natural habitat²⁰ forms the basis for Fig. 4a. The distorted image for the stated conditions is in Fig. 4b. Again many variations are possible in this distortion: the figure can become stockier or narrower; it need not have a thin 'waist'; and the apparently fierce 'mouth' and 'teeth' at the top of the figure can become much larger.

Example 3: A photograph of a present-day merman is given in Fig. 5b. It appeared over Lake Winnipeg on 2 May 1980, a hot calm day, while there was still some ice on the lake. The inland temperature at the time was about 28 °C, while the lake surface was near 0 °C. The source of the apparition, subsequently identified, was an insignificant boulder (Fig. 3a) sited 1.10 km from the camera.

Comparison with mediaeval descriptions

The author of the *King's Mirror* describes the merman with great care, while the less well-informed author of the *Historia Norvegiae* provides only a very general description. Yet the overall impression is identical.

Both accounts describe the merman as a huge animal rising straight out of the water. The lack of detail in the *Historia Norvegiae* implies that the animal had only been seen at some distance, a fact confirmed by direct statement in the *King's Mirror*. The enormous size could, therefore, only have been estimated by comparing its appearance with that of known objects at apparently similar distances.

The man-like shape implied by its name is consistent with the visual impression of an object viewed at some distance, when vertically elongated. As the horizontal dimension remains unchanged, the resulting image acquires a large height-to-width ratio, a form identified with man.

The accounts diverge only in the detailed configuration and in the specific features. The *Historia Norvegiae* suggests a cylindrical body, while the *King's Mirror* gives it shoulders like a man's, with a tapering body. As shown in Figs 2b, 4b and 5b, both shapes are a natural consequence of refractive distortion. Further, the *Historia Norvegiae* claims the merman has no head at all, while the *King's Mirror* lists a pointed head, having eyes, mouth, nose, chin and neck. Again, the process of refraction permits various upper terminations, and refracted markings in approximately the right place (Figs 2b, 4b) can create the impression of a face. As small atmospheric irregularities tend to disintegrate the image, producing a whiskery or hairy effect⁹, such characteristics could be attributed to its mate, the mermaid. At the distances implied in the *King's Mirror*, the human eye,

whose resolving power is between 0.5 and 1 arc min, registers only the major outlines; finer details are interpolated by the perception process. Eventually, numerous sightings similar in overall impression but differing somewhat in detail would result in a generalized description and the identification of a species.

The elevation of the observer is important in all optically induced anomaly sightings. The refractive distortion generally diminishes if the image can be viewed from above the thermocline. However, such variations would go unnoticed by Norse mariners, for the working of their ships was carried out from the deck, a few metres above the sea²¹. The subsequent use of much higher decked ships and the establishment of elevated lookouts would explain the infrequent merman sightings in later centuries.

Finally, the distorted boulder of Fig. 5b bears a striking resemblance to the *umibohzu*, or "priest of the sea", a giant-like monster known to the Japanese²². An even stronger likeness occurs in a narrow-necked version of Fig. 2b, in which the *umibohzu's* eyes are produced by the whale's white eye patch. Clearly the Japanese were interpreting the same visual phenomena in much the same way as did the Norse.

Correlation with the environment

The thermoclines that generate merman images are best created when a warm air mass slowly moves over significantly cooler surface air. In conditions of relative calm, the boundary between the two air masses will possess the flatness and uniformity necessary for well-defined (but distorted) images.

A typical situation is illustrated by the last stages of a warm front, when the warm-cool interface has almost descended to the surface. This has been experimentally verified by Wegener's observations of good superior mirages in the North Atlantic^{11,23}, which he later correlated with the arrival of warm fronts²⁴. However, neither he nor subsequent polar researchers have studied the fine structure of the temperature distributions that the lowest air layers must possess to produce merman images.

Scoresby²⁵, an experienced Arctic whaler, recorded numerous examples of anomalous refraction which strongly resemble the marvels of the *King's Mirror*: the elevated horizon, a dense appearance of the atmosphere, and the vertically elongated hulls of ships as if in a heeling position. Though Scoresby recognized some forms of refraction phenomena and correlated them with an easterly wind, he attributed others to vapours in the atmosphere, or confused them with ice-blink. He thus associated unusual visual effects with storms but did not identify

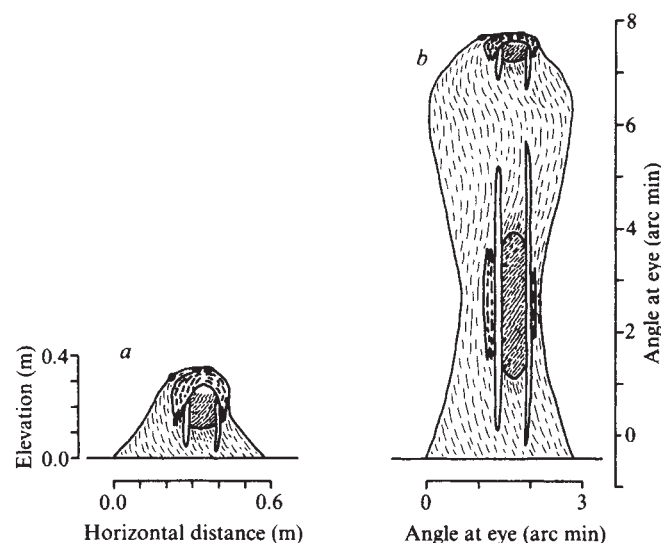


Fig. 4 *a*, Walrus. *b*, The distorted image of the walrus, seen from a distance of 0.7 km in the temperature conditions of Fig. 3a. Again, the *hafgerdingar* will be present if the same temperature profile exists over a sufficiently wide area. The horizon elevation would be +13', at a distance of 15 km, compared with -2' at 4.8 km for a normal atmosphere.

them with refraction. The dead calm followed by a sudden rise in temperature, the typical conditions noted by Scoresby just before a major storm, are ideally suited to the development of a thermocline. I have observed such an effect on a very localized scale, in the Canadian Arctic, over the Beaufort Sea: the sudden appearance of a medium-range mirage, followed within 40 min by a sudden temperature rise, rainstorms, and disappearance of the mirage. The amount of optical distortion depends directly on the temperature difference between the two air masses, which in turn determines the strength of the front and the severity of the subsequent storms.

Similar refractive distortions will have provided the Norse with sufficient opportunity to observe elongated images of the sea mammals which they used as navigational guides²⁶. Because Severin's experience in the *Brendan* indicates that whales do not avoid small sailing vessels²⁷, the limiting factor would have been the frequency of suitable atmospheric conditions rather than the presence of a suitable subject. These observations belonged to the thirteenth century mariner's general body of knowledge, and were recorded as such in the *King's Mirror*. It is fitting that the *King's Mirror* places both phenomena, the *hafgerdingar* and the merman, together in its text. From the optical reconstruction, as well as Scoresby's observations, clearly some form of *hafgerdingar* will be observed along with the merman. Both are correctly associated with the advent of storms on the open sea, and are located in the Greenland Sea where such extreme conditions were also observed by Scoresby.

Subsequent development

The precise descriptions in the *King's Mirror* directly contrast with those of Olaus Magnus three centuries later. His monsters, composites of several species assembled from fragmented reports¹⁹, reflect the general level to which Norse maritime knowledge had declined. As knowledge of the North Atlantic was regained by western European explorations, unrecognizable sightings reminiscent of the merman were occasionally reported^{28,29}. Nevertheless, the Norse merman had become largely a legendary creature, and skepticism as to its existence hardened when scientists identified it with the newly discovered dugong and manatee, both warm water mammals. To clarify this problem, Pontoppidan collected written accounts, as well as verbal reports from reliable witnesses. His attempt to isolate the factual from the mythical established the North European merman as an existing species with some well-defined characteristics. This study formed the basis for all future work even though Pontoppidan did not have access to the *King's Mirror*, a manuscript long believed lost¹.

It is of interest to examine the evolution of the thirteenth century merman into Pontoppidan's eighteenth century composite. Not only had the natural habitat of the merman shifted from the Greenland Sea to the coastal waters of northern Europe, but the conditions in which the merman was sighted had also changed. Many of the sightings accepted by Pontoppidan were reported on very warm, calm days. These observations suggest that the characteristic vertically distended appearance of an optically distorted image was still being identified as a merman; for such weather conditions, though different from those of the advancing front, also develop strong temperature gradients and consequent refraction. Further, while the specific correlation with frontal activity on the open sea had totally disappeared, the fear of the merman had survived as a generalized bad omen. Finally, the species had become approachable; whereas the early Norse reported only distant sightings, actual catches and findings of partially decomposed remains were claimed in the seventeenth and eighteenth centuries.

A similar evolution can be traced for the *hafgufa*. As originally described, it was a featureless monster of enormous dimensions in the Iceland Sea⁶. So few sightings had ever been reported that the author of the *King's Mirror* estimated its world population at one or perhaps two, and hesitated to include it in a compendium of sea life. By the eighteenth century, this monster had multiplied remarkably to inhabit the entire

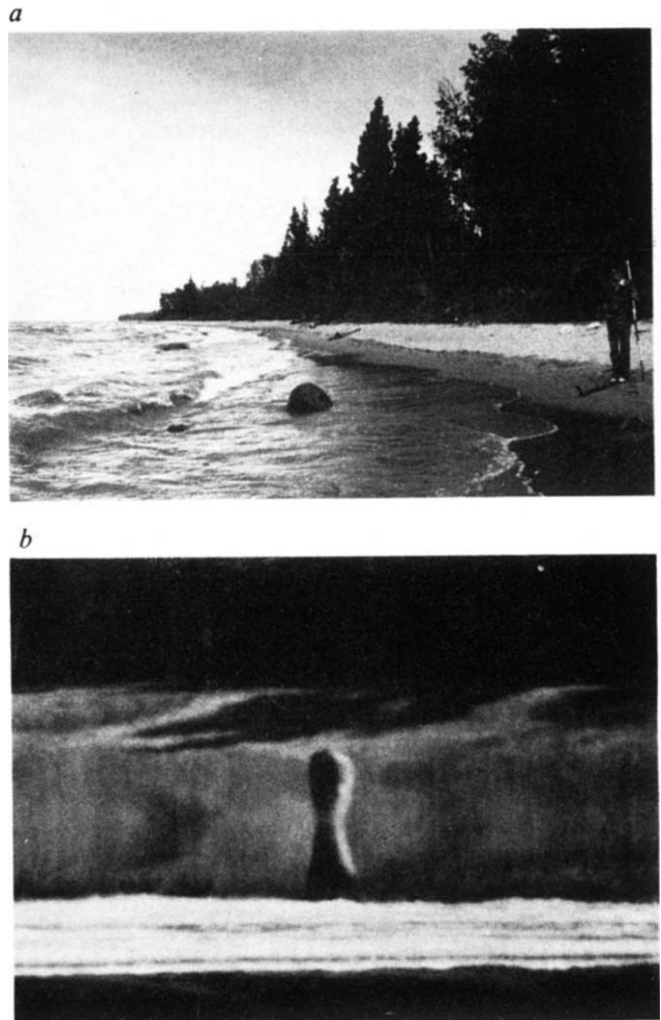


Fig. 5 *a*, Lake Winnipeg, 28 May 1980. The boulder in centre foreground was identified as the source of the image in Fig. 5*b*. It is ~68 cm wide and should show 30–35 cm above water on a calm day. *b*, The boulder distorted into a merman shape, 3:17 p.m. CDT (central daylight time), 2 May 1980. The image subtends 5.4' vertically and 2.2' horizontally. The photograph was made from a distance of 1.10 km using a mirror lens (focal length 1,250 mm) on Kodachrome 64 film. the exposure was 1/250 s at *f*/11, and the lens elevation was 2.5 m above the lake. A photograph taken just previously, at 3:15.5 p.m., shows the distortion as incompletely developed. The merman image of this boulder lasted for only a few minutes. Such conditions occur almost every spring on Lake Winnipeg.

North Atlantic, especially the coastal waters of Norway. Pontoppidan recorded the varied accounts of this creature but found the information so diverse and inconsistent that he did not attempt to identify it with any known species¹. However, by the end of the nineteenth century, the kraken had assumed the characteristics of a giant squid, an idea which has become increasingly accepted².

The *King's Mirror* is so accurate in its description of the merman and the *hafgerdingar* that a serious effort should be made to identify a natural phenomenon closely fitting the original *hafgufa*. Its appearance suggests that it may be a consequence of submarine volcanic activity in Icelandic waters. Steenstrup²⁶ used this theory to explain the *hafgerdingar*. But the *hafgerdingar* are much better described as an optical phenomenon, while Steenstrup's theory is a superior match, both in appearance and location, to the thirteenth century *hafgufa*.

Following the deterioration of contact between Iceland, Greenland and Norway, the merman and the *hafgufa* had become culturally isolated. Cut off from their origins, these

stories were free to develop according to local conditions. Progress in identifying these marvels was hampered by the fact that they had evolved in time, a problem similar to the one identified in a recent analysis of the Gunnbjorn Skerries³⁰. When renewed interest in the North Atlantic stimulated scholars to seek out all previously known information, both the skerries and the marvels of the *King's Mirror* emerged as composite remnants from the mediaeval age.

Conclusion

The merman described in the early sources, like the *hafgerdingar*, is, in fact, an accurate description of an unusual optical phenomenon. Its appearance is exactly that predicted by computer simulations of anomalous atmospheres, and the weather conditions with which it is associated agree with modern observations in arctic regions. The correspondence of these two marvels with reality leads directly to the conclusion that the

hafgufa account must also be considered as an observed natural phenomenon.

The extent to which Norse expertise deteriorated, once trans-Atlantic navigation ceased, is clearly demonstrated by the many unrealistic accounts found in sixteenth century sources. Clear and precise reports were superseded by confusion and the supernatural. The tendency of modern scholars has, therefore, been to dismiss as inaccurate or superstitious any observations inconsistent with their own experience. The optical interpretation of the merman and the *hafgerdingar* enables a clear differentiation between thirteenth and sixteenth century concepts and furthers an appreciation of the visual impressions which formed an integral part of Norse maritime experience.

This research was supported in part by the University of Manitoba's Research Board and Northern Studies Committee, as well as by NSERC grant A3196. We thank Dr F. W. Hougen for checking our Norwegian translation.

Received 21 July; accepted 14 November 1980.

1. Pontoppidan, E. *The Natural History of Norway*, II, 186–195, vi, 210–218 (Lind, London, 1755).
2. Oudemans, A. C. *The Great Sea Serpent* (Brill, Leiden, 1892).
3. Heuvelmans, B. *In the Wake of Sea Serpents* (Hill and Wang, New York, 1968).
4. Westrum, R. in *The Sociological Review Monograph No. 27* (ed. Wallis, R.) 293–314 (University of Keele, 1979).
5. Nansen, F. *In Northern Mists*, II, 243 (Heinemann, London, 1911).
6. Larson, L. M. (transl.) *The King's Mirror*, 135–138, 125 (Twayne, New York, 1917).
7. Lehn, W. H., Sawatzky, H. L. & Schroeder, I. *Scand. Rev.* **66**, 2, 36–41 (1978).
8. Koht, H. (transl.) *Historia Norvegiae (Den eldste Noregs-Historia)*, 13 (det Norske Samlaget, Oslo, 1921).
9. Lehn, W. H. *Science* **205**, 183–185 (1979).
10. Pernter, J. M. & Exner, F. *Meteorologische Optik* 2nd edn (Braumüller, Vienna, 1922).
11. Koch, J. P. & Wegener, A. *Meddr Grønland* **75**, 609–629 (1930).
12. Lehn, W. H. *J. opt. Soc. Am.* **69**, 776–781 (1979).
13. Lehn, W. H. & Sawatzky, H. L. *Polarforschung* **45**, 120–128 (1975).
14. Lehn, W. H. & El-Arini, M. B. *Appl. Opt.* **17**, 3146–3151 (1978).
15. Nickerson, R. *Brother Whale*, 101–102, 117 (Chronicle, San Francisco, 1977).
16. Slijper, E. J. *Whales and Dolphins*, 119 (University of Michigan Press, Ann Arbor, 1976).
17. Shackleton, E. H. *South—The Story of Shackleton's Last Expedition 1914–1917*, 22, 35, 323 (Macmillan, New York, 1920).
18. Norman, J. R. & Fraser, F. C. *Field Book of Giant Fishes*, 289–294 (Putnam, New York, 1949).
19. Matthews, L. H. *The Whale*, 248, 22–32 (Allen & Unwin, London, 1968).
20. Bruemmer, F. *Encounters with Arctic Animals*, 73 (McGraw-Hill Ryerson, Toronto, 1972).
21. Falk, H. S. *Altnordisches Seewesen*, 67 (Winter, Heidelberg, 1912).
22. Sweeney, J. B. *A Pictorial History of Sea Monsters*, 49 (Crown, New York, 1972).
23. Wegener, A. *Meddr Grønland* **42**, 125–356 (1914).
24. Wegener, A. *Annaln. Hydrogr. Berl. (Köppenheft)*, 93–95 (1926).
25. Scoresby, W. *An Account of the Arctic Regions, with a History and Description of the Northern Whale-Fishery*, I, 387–391, 442, 300, 398–399 (1820; reprinted by Kelley, New York, 1969).
26. Jones, G. *The Norse Atlantic Saga*, 115, 146 (Oxford University Press, London, 1964).
27. Severin, T. *The Brendan Voyage*, 124–132, 156, 188 (Arrow, London, 1979).
28. Best, G. *The Three Voyages of Martin Frobisher* (ed. Stefansson, V.) I, 100–101, Ivi (Argonaut, London, 1938).
29. Gould, R. T. *The Case for the Sea Serpent*, 12–16 (Philip Allan, London, 1930).
30. Lehn, W. H. & Schroeder, I. *Polarforschung* **49**, 173–187 (1979).

Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution

I. A. Wilson, J. J. Skehel* & D. C. Wiley

Gibbs Laboratory, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

*National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The haemagglutinin glycoprotein of influenza virus is a trimer comprising two structurally distinct regions: a triple-stranded coiled-coil of α -helices extends 76 Å from the membrane and a globular region of antiparallel β -sheet, which contains the receptor binding site and the variable antigenic determinants, is positioned on top of this stem. Each subunit has an unusual loop-like topology, starting at the membrane, extending 135 Å distally and folding back to enter the membrane.

INFLUENZA virus acquires a lipid membrane during maturation by 'budding' through the host-cell membrane. The resulting viral membrane contains the two virally coded integral membrane glycoproteins, haemagglutinin and neuraminidase¹. The haemagglutinin (HA) of the A/Hong Kong/1968 virus is a trimer of 224,640 molecular weight (MW). It may be solubilized from the viral membrane by bromelain digestion, which removes a 5,406 MW C-terminal hydrophobic (anchoring) peptide from each subunit^{2–4}. This observation, extended by the results of subsequent primary sequencing experiments^{5–14}, places the haemagglutinin in a class of integral membrane proteins characterized by a three-domain structure with a large hydrophilic, carbohydrate-containing domain on the external surface of the membrane, a small, uncharged hydrophobic peptide of 24–28 amino acids spanning the membrane, and a small, hydrophilic domain (10–55 amino acids) on the internal side of the membrane (Fig. 1a). Other well characterized examples of this class of membrane glycoprotein include the

human and murine histocompatibility antigens HLA and H-2 (refs 15, 16), erythrocyte glycoporphin¹⁷, and the animal virus membrane proteins. Such proteins show enzyme or 'recognition' activities, in contrast to proteins active in membrane transport, of which the structurally best defined example is bacterial rhodopsin^{18,19}.

Typical of membrane glycoproteins, the haemagglutinin chain is initially synthesized to include an N-terminal hydrophobic 'signal' peptide^{11,20,21}, which is subsequently removed as part of the process by which the protein is transported across and anchored into the membrane^{22,23} (Fig. 1a). Each polypeptide chain of the trimer is glycosylated at seven sites with a total carbohydrate of 13,000 MW (19% by weight)^{6,11}. We wish to understand the structural correlates of these properties. What is the nature of the membrane attachment of this protein? What is the location and biological role of the carbohydrate? As a virus membrane protein, the haemagglutinin expresses a number of surface activities whose structural bases are also sought. How

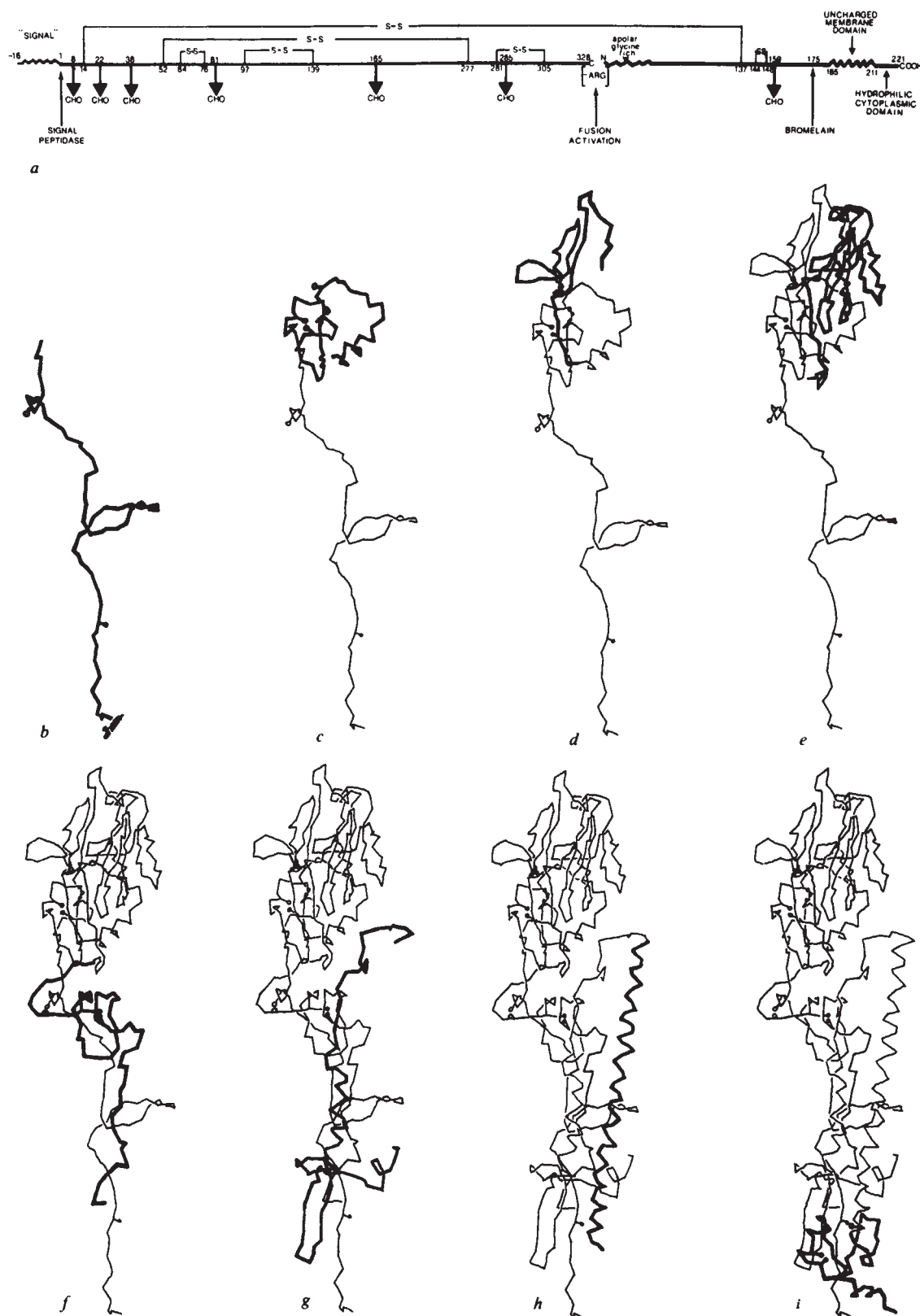


Fig. 1 a, Schematic diagram of the 1968 Hong Kong haemagglutinin sequence showing the external domain (ends 185 HA₂), membrane anchoring domain (185–211) and cytoplasmic domain (212–221). The signal sequence, S-S bridges, CHO-attachment sites and fusion activation site (see text) are also shown. b–i, Schematic diagrams showing the path of the polypeptide chain of a monomer (HA₁ and HA₂) traced from the N-terminus of HA₁ to the C-terminus of HA₂. d, e, Ignoring loop-out regions and focusing on the eight extended β -strands, the 'Swiss roll' is easily modelled with a strip of paper. Using only one side of the paper, draw two antiparallel lines (label with arrows) along the length of the strip, one on the left and one on the right. Number the right line from 1 to 4 evenly from the bottom of the line to the top. Connect the top of the right line to the top of the left line and number the left line 5–8 from top to bottom. The numbers represent the eight strands of the antiparallel β -structure—note that 4 is antiparallel to 5, etc. Fold the strip of paper into four equal lengths (line on outside) and arrange the resulting 'Swiss roll' so that strands 1, 8, 3 and 6 are visible in that order from one side and 2, 7, 4 and 5 are seen from the other (Fig. 2). The point between strands 4 and 5 corresponds to amino acid 208 (HA₁) in the haemagglutinin structure (Fig. 3a). The eight strands are approximately 1, 119–123; 8, 251–259; 3, 174–182; 6, 229–237; 2, 163–170; 7, 241–248; 4, 200–205 and 5, 210–213 (Fig. 3a). (This figure is intended as a description of the structure, not a folding pathway.)

does this protein recognize the sialic acid-containing host-cell receptor²⁴? How does the post-translational removal of an arginine (329), dividing the chain into two disulphide-linked chains, HA₁ (328 amino acids) and HA₂ (221 amino acids)^{11,14}, activate infectivity and membrane fusion activity²⁵⁻²⁷? How does variation in the amino acid sequence of the structure alter the antigenic determinants of the molecule and lead to recurrent epidemics of respiratory disease?

We report here a first description of the results of the 3 Å-structure determination of the external, glycoprotein domain of the influenza virus haemagglutinin, which provides initial answers to certain of the above questions. The host-receptor binding site and antigenic determinants are located on a globular region which lies on top of a long fibrous coiled-coil (Fig. 2). The fusion activation peptide is located near (35 Å) the virus membrane end of the molecule. The polypeptide fold is unusual, like a long loop, starting at the membrane, extending 135 Å distally, and ending at the membrane. This tertiary structure indicates that the molecule is probably extruded from the membrane as a single chain during biosynthesis and may even remain attached via the signal peptide to the membrane during folding. Most of the oligosaccharide chains are found near the membrane end of the molecule. The identification of antibody binding sites and the contribution of their variation to the recurrence of respiratory disease in man is discussed in the accompanying article²⁸.

Structure determination

The bromelain-released haemagglutinin (30 mg ml⁻¹) crystallizes in tetragonal space group P4₁, $a = 163.2$ Å, $c = 177.4$ Å from 1.28 M sodium citrate, 0.1% azide, pH 7.5 (ref. 29). A single isomorphous derivative, mercury phenylglyoxal³⁰, was used in structure determination. 1° oscillation photographs (12-h exposure) were recorded at 4°C, using an Elliot G-X-6 rotating anode source operated with a 100 µm focus at 40 kV

Table 1 Data collection statistics

Data set	No. films	Total reflections measured	Unique reflections	R_{sym}	R_{whole}	R_{asym}
Native	78	408,672	83,015	0.12	0.11	0.18
Derivative	62	338,738	78,077	0.13	0.13	0.20

$R = \sum |s_i I_{hi} - \bar{I}| / \sum \bar{I}_n$ where s_i = scale factor for crystal i , I_{hi} = intensity of reflection h of crystal i and $\bar{I}_n$ = average intensity of reflections. R_{sym} , symmetry-related reflections on the same film; R_{whole} , fully recorded reflections from different films; R_{asym} , all reflections ($\geq 50\%$ recorded) on different films, final scaling.

and 20 mA, and Franks focusing optics³¹. Up to 15° of data could be recorded from each crystal by exposing a fresh volume to X rays after 24 h exposure (Table 1).

Solution of the 3 Å difference Patterson revealed six heavy-atom locations per molecule (Table 2a), two sites 27 Å apart per monomer, which define a non-crystallographic, molecular 3-fold axis. (Details of the structure determination will be published elsewhere.) The asymmetric unit contains one glycoprotein trimer³² of 208,422 MW and is 78% solvent (by weight). The standard isomorphous replacement (SIR) phases of 59,669 reflections were determined to 3 Å resolution with an overall figure of merit of 0.34 and overall phasing power of 1.38 (Table 2b). Eleven cycles of refinement of these and previously undetermined phases of a total of 74,194 reflections by non-crystallographic 3-fold symmetry averaging³³ dramatically improved the quality of the electron density maps and the agreement of calculated and observed amplitudes ($R = 0.47$ cycle 1, $R = 0.27$ cycle 11). The molecular envelope required redetermination three times during refinement to include protruding regions of protein and carbohydrate that had been truncated by the original estimate of the envelope but that were clearly apparent in difference Fourier maps.

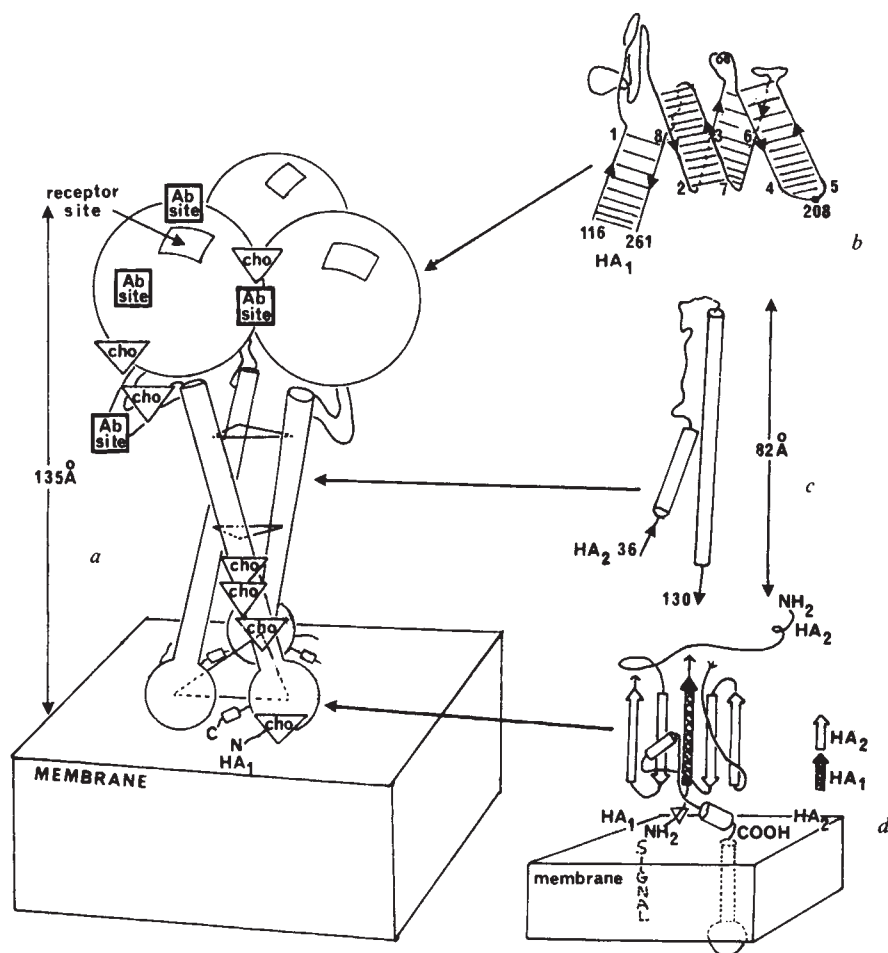


Fig. 2 *a*, Schematic diagram of the 1968 haemagglutinin trimer showing carbohydrate attachment sites (CHO), antigenic sites (Ab site) (see ref. 28), the host-receptor binding site and the arrangement relative to the membrane. *b*, The eight-stranded β -sheet structure and looped-out region in the globular domain. *c*, HA₂ residues 36–130, including two α -helices (cylinders), form part of the fibrous region. Three of the long helices, one from each monomer, pack together as a triple-stranded coiled-coil which stabilizes the trimer. *d*, The membrane end of the molecule contains a five-stranded β -sheet. The central strand is the N-terminal of HA₁ and the adjacent strands come from the C-terminal chain of HA₂. Broken lines suggest the path of the hydrophobic anchoring peptide cleaved by bromelain. The possible attachment of the N-terminus to the membrane by the signal sequence is also indicated. The site of the first oligosaccharide chain (residue 8) is shown as a triangle.

Interpretation of electron density map

A minimap ($4 \text{ \AA cm}^{-1}$) was calculated for the trimer, sectioned normal to the 3-fold axis. The subunit boundaries were clearly visible. Ambiguities in chain tracing were resolved with the aid of the amino acid sequence (provided by Fiers *et al.*¹¹ and Ward *et al.*¹⁴). α -Carbon positions of residues 4–328 of HA₁ and residues 1–175 of HA₂ have been located. The amino terminus of HA₁ and the C-terminus of HA₂ are close to the membrane end of the haemagglutinin and therefore to the bromelain cleavage site at residue 175 of HA₂ (ref. 34). The unaveraged map indicates that deviations from 3-fold symmetry may occur in the amino-terminal region of HA₁ by interaction of one of the amino-termini with an adjacent molecule in the crystal. The side-chain densities are well defined and unambiguous for 50–60% of the residues and another 20–30% fit moderately well. Only 15% of the residues have weak side-chain or main-chain density, and the chain can clearly be traced throughout the entire molecule.

Location of the six disulphides showed that only one disulphide linkage connects HA₁ and HA₂ (refs 7, 35).

Description of the structure

The haemagglutinin trimer is an elongated cylinder, 135 Å long, with a triangular cross-section varying in radius from 15 to 40 Å. The structure is divided into two distinct regions: (1) a long fibrous region, containing residues from both HA₂ and HA₁, is centred on a 76 Å triple-stranded α -helical coiled-coil extending from the membrane surface; (2) a globular region containing residues entirely from HA₁ sits on top of the fibrous stem and includes an 8-stranded β -structure.

Assembly of the trimer seems to be a requirement for stabilization of the elements of the subunit tertiary structure in the fibrous stem region.

Tertiary fold. The unusually extended tertiary structure of the haemagglutinin may be appreciated by following the schematic drawing of its chain tracing (Fig. 1). The first 63 amino acids reach from the N-terminus at the membrane end of the molecule in an almost extended conformation 96 Å along the molecular length before becoming involved in a compact fold (Fig. 1*b, c*). The fold (Fig. 1*c*) contains a short disulphide loop (64–76) and a three-turn helix.

β -Structure. The distal end of the molecule forms an 8-stranded antiparallel β -sheet structure from residue 116 to 261 (Fig. 1*d, e*; see also Fig. 3). Similar 'Swiss roll' β -structures have been described in tomato bushy stunt virus³⁶ and southern bean mosaic virus³⁷, and have been discussed by Richardson³⁸. Two looped-out regions between strands in the β -structure seem to be important antigenic regions: the 125–163 loop between β -strands 1 and 2, and the 187–199 α -helix-containing loop between β -strands 3 and 4 (ref. 28). This β -structure also provides the framework for the sialic acid-receptor binding site (see below).

Switch region. The globular β -region reconnects to the fibrous stem by only two disulphide loops 52–277 and 281–305 (Fig. 1*f*). The globular region is thus an HA₁ loop connected to the fibrous stem by only two antiparallel chains, the N-terminal chain extending away from the membrane and the C-terminal chain returning towards the membrane. In the fibrous region the C-terminal chain of HA₁ extends 60 Å towards the membrane end of the molecule, remaining antiparallel to the amino-terminal HA₁ chain.

HA₁–HA₂ junction. The N-terminus of HA₂ is 21 Å from the C-terminus of HA₁, indicating a substantial rearrangement in this region when the haemagglutinin is activated by cleavage (Fig. 1*g*). The HA₂ amino-terminal glycine-rich sequence forms an uncommon series of four contiguous reverse turns before wrapping transversely around the fibrous coil region ~35 Å from the membrane end of the molecule.

Fibrous helices. Two antiparallel α -helices form the backbone of the fibrous region of each subunit. A short α -helix (29 Å) extends distally from the membrane and connects by an extended chain to a very long (53 amino acid, 14 turn) α -helix which stretches 76 Å back towards the membrane (Fig. 1*g, h*). In the trimer, this long helix twists around two identical helices from other subunits to form the coiled-coil core of the fibrous region. The conformation of the extended chain (HA₂, 59–76) and the first half of the long α -helix (HA₂, 77–100) seems to be stabilized predominantly by extensive contacts with the other subunits of the trimer (Fig. 3).

Terminal globular region. The membrane end of the fibrous region contains a compact globular fold which includes a 5-stranded antiparallel β -sheet. The structure of this sheet is noteworthy—the central strand (no. 3) is the amino-terminal chain of HA₁, and the extended chains of HA₂, 350 amino acids away in the sequence, form the two strands flanking it on each side (nos 1, 2 and 4, 5) (Figs 1*g–i* and 2). The chain terminates in a three-turn α -helix at the bromelain cleavage site, Gly 175 (ref. 34), at the extreme membrane end of the molecule (Fig. 1*i*). An uncleaved haemagglutinin would enter the membrane 10 amino acids further along. At this stage in the analysis we cannot rule out the possibility that the C-terminal helix was normal to the membrane and may have rearranged as a result of bromelain digestion.

The complete HA₁ chain contains about 36% β -sheet, 6% α -helix and 16% reverse turns. The HA₂ chain has a significantly more regular secondary structure with 14% β -sheet, 53% α -helix and 15% reverse turns.

Biosynthesis and folding. The preceding description of the tertiary structure is not intended to suggest a pathway for the folding of the protein. The acquisition of the three-dimensional fold of a molecule that has been synthesized in the cytoplasm and that has had most of its mass transported to the external side of the membrane is poorly understood.

The tertiary fold of the haemagglutinin structure suggests that the molecule was extruded from the membrane as a single chain

Table 2 Solution of 3 Å difference Patterson with SIR phase refined parameters and statistics

<i>a</i>		Atom site	<i>x</i>	<i>y</i>	<i>z</i> (Å)	Occupancy	<i>B</i> (Å ²)			
		1	30.9	53.4	0.0*	0.47	18.6			
		2	25.1	54.0	21.9	0.60	13.4			
		3	0.4	43.0	−16.5	0.36	16.4			
		4	3.8	57.7	38.6	0.38	20.8			
		5	13.0	41.6	7.3	0.47	19.2			
		6	41.7	78.5	1.0	0.35	17.4			
	<i>b</i>	Resolution (Å)	8.7	6.9	5.7	4.8	4.2	3.7	3.3	3.0
No. reflections		989	2,223	3,968	5,726	8,206	11,431	13,859	13,267	59,669
<i>M</i>		0.40	0.41	0.38	0.36	0.36	0.34	0.33	0.31	0.34
<i>F_c/E_i</i>		1.54	1.74	1.69	1.58	1.47	1.39	1.27	1.10	1.38
<i>R</i> _{KRAUT}		0.11	0.13	0.14	0.13	0.13	0.15	0.18	0.24	0.16

Occupancy, relative fractional occupancies (not absolute). B, isotropic temperature factor.

* Atom 1 z fixed at 0.0.

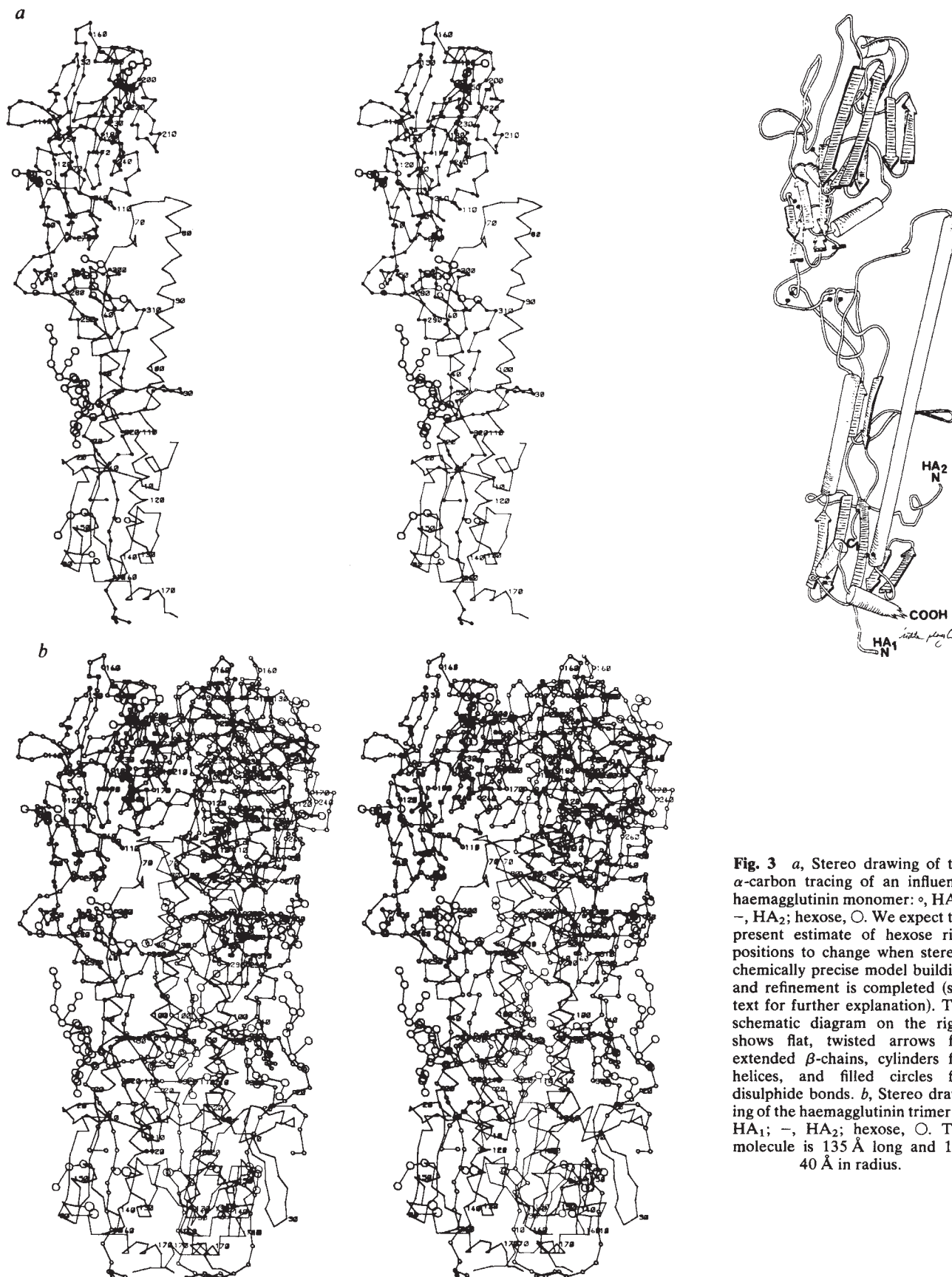


Fig. 3 *a*, Stereo drawing of the α -carbon tracing of an influenza haemagglutinin monomer: $\circ$, HA₁; $\triangle$, HA₂; hexose, $\circ$. We expect the present estimate of hexose ring positions to change when stereochemically precise model building and refinement is completed (see text for further explanation). The schematic diagram on the right shows flat, twisted arrows for extended β -chains, cylinders for helices, and filled circles for disulphide bonds. *b*, Stereo drawing of the haemagglutinin trimer: $\circ$, HA₁; $\triangle$, HA₂; hexose, $\circ$. The molecule is 135 Å long and 15–40 Å in radius.

(a globular region found at the end of a 'structureless' first 63 amino acids) as opposed to a more cooperative refolding that can be imagined if the entire molecule emerged from the membrane by a concerted conformational change^{23,39}.

Furthermore, the loop-like topology, with both the N- and C-termini of the chain at the extreme membrane end of the

molecule, suggests that the structure may have folded, attached to the membrane not only by the carboxy-terminal anchor peptide (or extruding chain) but also by the amino-terminal hydrophobic signal sequence. (In this respect, the shortest uncharged section in all known haemagglutinin signal sequences (A/Memphis/10/78[H1N1])⁴⁰ is 11 amino acids long, sufficient

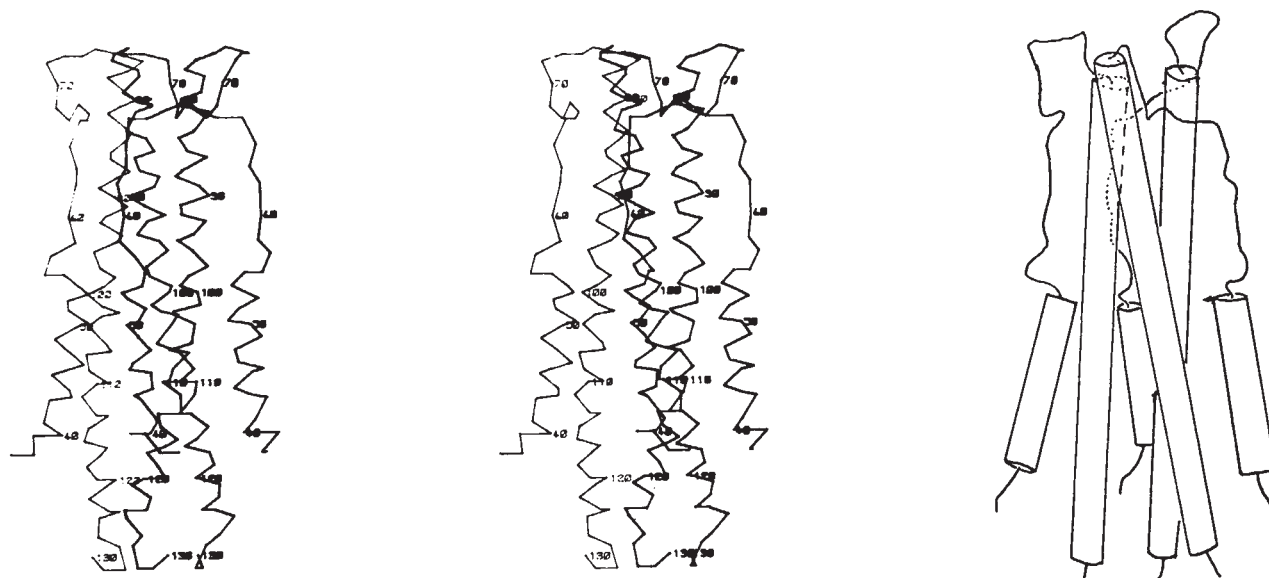


Fig. 4 Stereo drawing of amino acids 36–130 of the HA₂ chain, illustrating the triple-stranded coiled-coil core of the trimer. The schematic drawing on the right illustrates the same region with cylinders representing helices.

to span the lipid bilayer in an extended conformation.) Although the peculiar tertiary structure of the haemagglutinin could be generated without either of these suggestions, they emerge naturally from a consideration of the molecular form.

Quaternary structure

Coiled-coil. The major forces stabilizing the haemagglutinin's trimeric subunit interactions arise from a triple-stranded coiled-coil^{41,42} in the fibrous region of the molecule. An α -helix (74–126) extends 76 Å along the molecular 3-fold symmetry axis. Three such helices, one from each subunit, pack together and twist 100° around each other in a left-handed superhelix (Fig. 4). The N-terminal (top) half of the coiled-coil is tightly packed (10 Å helix to helix) with many nonpolar residues in van der Waals contact around the 3-fold axis (Ile 77, Leu 80, Val 84, Leu 91, Leu 98, 99, 102). The C-terminal half of the coiled-coil is expanded away from the 3-fold axis (up to 22 Å helix to helix), with polar and charged residues interacting around the 3-fold axis (His 106, Asp 109, Asp 112, Ser 113, Asn 116, Glu 120, Arg 123). Some polar interactions may be mediated by solvent molecules.

Numerous buried charge pairs form intra-chain and inter-subunit salt bridges that contribute to the stability of the coil. The amino-terminal residues Leu and Phe of HA₂ pack around the 3-fold symmetry axis in the beginning of the open region of the coil (112–118).

The coiled-coil interactions agree with theoretical predictions of helix packing residues based on a heptad repeat⁴³, but the a, d contact region drifts through a, e, to b, e, by the end of the helices. Ward and Dopheide⁴⁴ predicted a coiled-coil in the approximately correct location from heptad repeats in their haemagglutinin sequence data.

Three short α -helices (HA₂, 38–56) pack almost hexagonally and antiparallel with the lower half of the three long helices (Fig. 4). The interactions again involve a twisting such that one end of the short helix lies between two long helices and the other end lies against the long helix of its own subunit.

Trimer assembly. Assembly of the trimeric 'spike' provides stabilizing interactions which seem to be crucial for the tertiary folding of regions in the fibrous stem of the individual subunits. The top half of the long α -helix (HA₂, 77–100) has very few interactions within its own subunit and in a monomer would expose a 36 Å hydrophobic exterior surface. Similarly, the extended chain, HA₂ 65–77, and the β -loop of HA₁, amino acid 30, make most of their stabilizing contacts with an adjacent subunit. Thus the polypeptide chain folding in these regions of the monomer may only be completed on assembly of monomers

to form the trimers. A corollary of these observations is that the molecule may only project off the surface of the membrane as a 'spike' when assembled as a trimer. Whether the possible flexibility in the monomer stem region in the absence of trimeric contacts may be exploited (by assembly–disassembly) for any biological role remains to be tested.

Membrane-associated region. The top half of the triple-stranded coiled-coil in the fibrous region of the haemagglutinin suggests a model for the quaternary interactions that may take place in the uncharged, predominantly hydrophobic, C-terminal peptide of the complete haemagglutinin (Fig. 1a), the domain that anchors the molecule into the virus membrane. This membrane-associated peptide (185–211) is probably α -helical. Twenty-seven uncharged amino acids are sufficient to cross the lipid bilayer in a helical conformation and the hydrogen bonding along the helix would prevent the unfavourable occurrence of unpaired carbonyl oxygen and amide nitrogen atoms in the lipid environment. Whether the three chains of the haemagglutinin trimer cross the membrane together as a unit or separately is unknown. A triple-stranded coiled-coil of α -helix crossing the membrane as a unit is suggested as it would facilitate trimerization on the internal side of the membrane, a property which may be important in the mechanism of budding or other transmembrane functions.

Globular region trimer contacts. The quaternary interactions of the globular region are far fewer than in the fibrous region. Residues from the eight-stranded β -structure are in contact but the interface is not extensive (Fig. 3b). The oligosaccharide attached at residue 165 (HA₁) spans the interface between the subunits, forming substantial contacts on both sides that would stabilize the oligomer in this region.

Carbohydrate

The carbohydrate (MW 13,000 per monomer) is located on the surface of the trimer, all along its length, with more in the half of the molecule nearest the membrane surface than at the distal end (Fig. 3). Seven Asn-X-Thr/Ser (six in HA₁) sequences occur in the Hong Kong haemagglutinin and all are glycosylated^{6,11}. Two of the oligosaccharide moieties are simple and contain only *N*-acetyl glucosamine and mannose (HA₁, 165, 285) whereas the others also contain galactose and fucose (HA₁, 8, 22, 38, and HA₂, 154). The absence of terminal sialic acids is due to the activity of the viral neuraminidase. The carbohydrate on residues HA₁ 165 and 285 is clearly visible in the electron density maps, whereas residues 22 and 38 (HA₁) have less well defined carbohydrate density and residues HA₁ 8, 81 and HA₂ 154 have

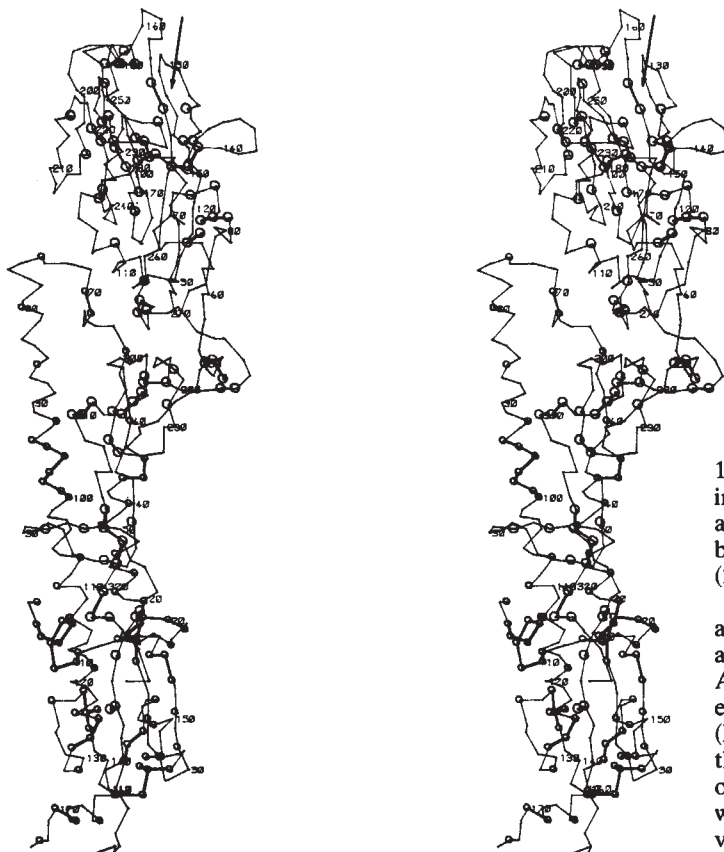


Fig. 5 Conserved amino acid positions from the known haemagglutinin sequences: ○, HA₁; ◐, HA₂. The pocket of highly conserved amino acids at the distal (top) end of the molecule, tentatively identified as the host-cell oligosaccharide receptor binding site, is marked by an arrow.

little or no visible carbohydrate in present maps, possibly due to positional disorder.

We have assigned preliminary positions to 63% of the hexose rings of the oligosaccharides. Our analysis is at present incomplete; for example, our placement of hexose rings is inconsistent with expected branch points at some locations (Fig. 3). Based on the experiences of Huber and colleagues with an oligosaccharide chain on IgG, we expect that when stereochemically precise model building and refinement are completed, these structures will be reliably determined^{45,46}.

The carbohydrate visible in the map covers 17–20% of the available surface and external residues of the trimer. Residues covered by the carbohydrate are mainly hydrophilic (60%), with Ser and Thr (33%) most frequently interacting with the sugars. Hydrophobic amino acids are often covered (33%) but no clusters of hydrophobic amino acids occur around any substituted Asn residue.

The role of carbohydrate in membrane antigens is largely undefined. Oligosaccharide chains on the haemagglutinin seem to have some structural roles, an observation also made for an IgG oligosaccharide⁴⁵. Carbohydrate–protein interactions are involved in the subunit–subunit interface and inter- and intra-chain contact regions. Thus these interactions may stabilize the monomer and the trimeric structure. We cannot assess the possibility that contacts between trimers on the viral surface could involve oligosaccharide chains. Two lysines, three arginines and five aromatic residues, potential tryptic and chymotryptic sites, are covered by oligosaccharide chains which suggests that carbohydrates may also stabilize the haemagglutinin by providing protection against degradative enzymes. The possibility that carbohydrate modulates recognition of the haemagglutinin by the immune system is discussed in the accompanying paper²⁸.

Structure comparison

Other influenza haemagglutinins seem to have structures very similar to that which we have found for the 1968 Hong Kong strain. The haemagglutinins from 12 strains of influenza have been sequenced (some partially) including viruses from 1957 to

1978 and an avian influenza^{5–14,47}. Sequence comparison indicates that 22% of HA₁ and 45% of HA₂ residues are absolutely conserved—these residues include the six disulphide bridges and many structurally important residues such as proline (21 residues) and glycine (9 residues).

Furthermore, the structure reveals that all known insertions and deletions in the sequences of these different strains can be accommodated at the surface of the 1968 Hong Kong structure. A core of structurally important conserved residues is also evident in both the globular and fibrous regions of the molecule (Fig. 5). Similarly, all the carbohydrate attachment sites found in the known sequences are on the surface and could accommodate oligosaccharide side chains, and detailed structures associated with antigenicity also seem to be conserved²⁸. These observations indicate that the 1968 structure is a good first-order model for haemagglutinins from many strains. Deviations from this structure will be important, however, in understanding the antigenic variation which the virus undergoes.

Biological activities

The distribution of conserved amino acids on the 3 Å structure reveals a highly conserved region in a surface pocket on the distal end of the molecule that seems well suited for a binding to host-cell oligosaccharide (Fig. 5). This tentative receptor binding pocket includes conserved residues Tyr 98, His 193, Glu 190, Trp 153 and Leu 194 (Fig. 5). It is similar to the wheat-germ agglutinin sialic acid-binding site⁴⁸. No direct experiment has associated virus–cell receptor binding with specific amino acids in the haemagglutinin. Crystallographic data on the haemagglutinin in the presence of oligosaccharides are being collected to locate the receptor binding site directly.

The most highly conserved sequence in the haemagglutinin is the amino terminus of HA₂, where 1 substitution in the first 11 and only 5 in the first 23 amino acids have been observed. This sequence is associated with the activity by which the virus penetrates a host-cell membrane to initiate infection. Cleavage into HA₁ and HA₂ at this point is required for infectious penetration and for *in vitro* membrane fusion^{26,27}. An homologous sequence is present at the cleavage-activation site on the Sendai virus fusion protein, where cleavage is required both for infectivity and membrane fusion activity⁴⁹. The HA₂ N-terminal sequence is strongly nonpolar, the first charged side chain occurring at position 11 (Glu). It is unexpectedly rich in glycines (1,4,8 (ref. 12) and 13,16,20,23 (refs 4, 13, 14)) which could indicate flexibility or the presence of an unusual conformation. In the present structure, the first 11 amino acids form a contiguous series of sharp bends, with residues 2 and 3 buried, making 3-fold contacts at the centre of the molecule (Figs 3b, 5). (The series of glycine-rich bends in our unrefined model have not been identified with any known form of helix.) The α-amino group forms a salt bridge to an aspartic acid. The strong conservation of the apolar sequence is particularly striking when compared with signal or C-terminal anchoring hydrophobic sequences. In both the latter cases, almost every position has accepted conservation apolar substitutions

(100% in signals, 85% in hydrophobic tails), which suggests a rather specific role for the terminal sequence in the fusion process.

There are two obvious difficulties in understanding the role of haemagglutinins in membrane fusion. First, the 135 Å length of the haemagglutinin seems to restrict membranes from any closer initial approach. Second, the HA₂ amino-terminus is ~100 Å from the distal tip and 35 Å from the membrane end of the molecule. For this region of the apolar terminal residues to approach either the virus or host-cell membrane a major conformational change would be required after receptor binding. Alternatively, some other region of the molecule significantly affected by the activation cleavage at the HA₂ amino terminus may participate.

Received 20 October; accepted 16 November 1980.

1. Laver, W. G. & Kilbourne, E. D. *Virology* **30**, 493–501 (1966).
2. Brand, C. M. & Skehel, J. J. *Nature new Biol.* **238**, 145–147 (1972).
3. Wiley, D. C. & Skehel, J. J. *Topics infect. Dis.* **3**, 135–138 (1978).
4. Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93–97 (1975).
5. Porter, A. G. *et al.* *Nature* **282**, 471–477 (1979).
6. Ward, C. W. & Doppeide, T. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 27–38 (Elsevier, Amsterdam, 1980).
7. Doppeide, T. A. & Ward, C. W. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 21–26 (Elsevier, Amsterdam, 1980).
8. Sleight, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 69–78 (Elsevier, Amsterdam, 1980).
9. Gething, M. J., Bye, J., Skehel, J. & Waterfield, M. D. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 1–10 (Elsevier, Amsterdam, 1980).
10. Min Jou, W. *et al.* in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 63–68 (Elsevier, Amsterdam, 1980).
11. Verhoeyen, M. *et al.* *Nature* **286**, 771–776 (1980).
12. Threlfall, G., Barber, C., Carey, N. & Emtage, S. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 51–61 (Elsevier, Amsterdam, 1980).
13. Min Jou, W. *et al.* *Cell* **19**, 683–696 (1980).
14. Ward, C. W. & Doppeide, T. A. *Biochem. J.* (in the press).
15. Springer, T. A. & Strominger, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2481–2485 (1976).
16. Coligan, J. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 3390–3394 (1978).
17. Tomita, M. & Marchesi, V. T. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2964–2968 (1975).
18. Henderson, R. & Unwin, P. N. T. *Nature* **257**, 28–32 (1975).
19. Engelman, D. M., Henderson, R., McLachlan, A. D. & Wallace, B. A. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2023–2027 (1980).
20. Air, G. M. *Virology* **97**, 468–472 (1979).
21. McCauley, J. *et al.* *FEBS Lett.* **108**, 422–426 (1979).
22. Milstein, C., Brownlee, G. G., Harrison, T. M. & Mathews, M. B. *Nature new Biol.* **239**, 117–120 (1972).

Finally, antibodies that neutralize viral infectivity interact with the haemagglutinin. Identification of antibody binding sites and discussion of how their variation contributes to the recurrence of influenza disease in man is reported in the accompanying paper²⁸.

We thank Colin Ward and Theo Doppeide for supplying results before publication, and Walter Fiers and colleagues for the sequence of the X:31 (AICHI/68) haemagglutinin. We acknowledge support from NIH AI13654 (DCW) and NSF PC-77-11398 (computing hardware). We thank Edward Gordon, David Stevens, Frank Escobar and Craig Steele for technical assistance, Jane Ladner for developing the stereo drawing program used for illustrations, and Hidde Ploegh for the schematic diagram in Fig. 3.

23. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 835–851 (1975).
24. Hirst, G. K. *J. exp. Med.* **75**, 49–64 (1942).
25. Lazarowitz, S. G. & Choppin, P. W. *Virology* **68**, 440–454 (1975).
26. Klenk, H. D., Rott, R., Orlich, M. & Blodorn, J. *Virology* **68**, 426–439 (1975).
27. Huang, R. T. C., Wahn, K., Klenk, H. D. & Rott, R. *Virology* **104**, 294–302 (1980).
28. Wiley, D. C., Wilson, I. A. & Skehel, J. J. *Nature* **289**, 373–378 (1981).
29. Wiley, D. C. & Skehel, J. J. *J. molec. Biol.* **112**, 343–347 (1977).
30. Monaco, H. L. thesis, Harvard Univ. (1978).
31. Harrison, S. C. *J. appl. Crystallogr.* **1**, 84–89 (1968).
32. Wiley, D. C., Skehel, J. J. & Waterfield, M. D. *Virology* **79**, 446–448 (1977).
33. Bricogne, G. *Acta crystallogr.* **A32**, 832–847 (1976).
34. Doppeide, T. A. & Ward, C. W. *J. gen. Virol.* (in the press).
35. Waterfield, M. D., Scrase, G. & Skehel, J. J. *Nature* **289**, 422–424 (1981).
36. Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* **276**, 368–373 (1978).
37. Abad-Zapatero, C. *et al.* *Nature* **286**, 33–39 (1980).
38. Richardson, J. *Adv. Protein Chem.* (in the press).
39. Wickner, W. A. *Rev. Biochem.* **48**, 23–45 (1979).
40. Air, G. M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 135–146 (Elsevier, Amsterdam, 1980).
41. Pauling, L. & Corey, R. B. *Nature* **171**, 59–61 (1953).
42. Cohen, C. & Holmes, K. C. *J. molec. Biol.* **6**, 423–432 (1963).
43. McLachlan, A. D. & Stewart, M. J. *J. molec. Biol.* **98**, 293–304 (1975).
44. Ward, C. W. & Doppeide, T. A. *Aust. J. biol. Sci.* **33**, 449–455 (1980).
45. Deisenhofer, J., Coleman, P. M., Epp, O. & Huber, R. *Hoppe-Seyler's Z. physiol. Chem.* **357**, 1421–1434 (1976).
46. Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. in *Proc. 7th Aharon Kelzir-Katchalsky Conf. Ginosar* (ed. Balasan, N.) (in the press).
47. Laver, W. G., Air, G. M., Doppeide, T. A. & Ward, C. W. *Nature* **283**, 454–457 (1980).
48. Wright, C. S. *J. molec. Biol.* **141**, 267–291 (1980).
49. Gething, M. J., White, J. M. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2737–2740 (1978).

Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation

D. C. Wiley & I. A. Wilson

Gibbs Laboratory, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138

J. J. Skehel

National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

Four 'antigenic sites' on the three-dimensional structure of the influenza haemagglutinin are identified. At least one amino acid substitution in each site seems to be required for the production of new epidemic strains between 1968 and 1975.

THE three-dimensional structure of the major surface antigen, the haemagglutinin, of influenza virus was determined to advance our understanding of how a membrane antigen is recognized by the immune system¹. Variation in the haemagglutinin's antigenic structure is associated with the familiar and uncontrolled recurrence of influenza epidemics in man². The structural basis for this phenomenon is sought to answer the following questions: which regions of the molecule are immunogenic and dominant in humoral and cellular immune recognition? How do the observed mutations result in an escape from neutralization? Why are conserved surface regions of the haemagglutinin, such as those involved in its many activities¹, not involved in the induction of cross-reacting antibodies? What

are the requirements for the emergence of a novel epidemic strain?

This report presents an analysis of the three-dimensional structure of the 1968 Hong Kong haemagglutinin using the available primary sequence information on natural and laboratory-selected antigenic variants^{3–16}. We propose initial answers to some of the above questions. Four antigenic sites are identified on the surface of the molecule. At least one amino acid substitution has been observed in each site of the haemagglutinins of recent epidemic strains and thus seems to be required for a new epidemic to occur. The probable stereochemical consequences of mutations at these sites and suggestions of how they might reduce antibody binding are discussed.

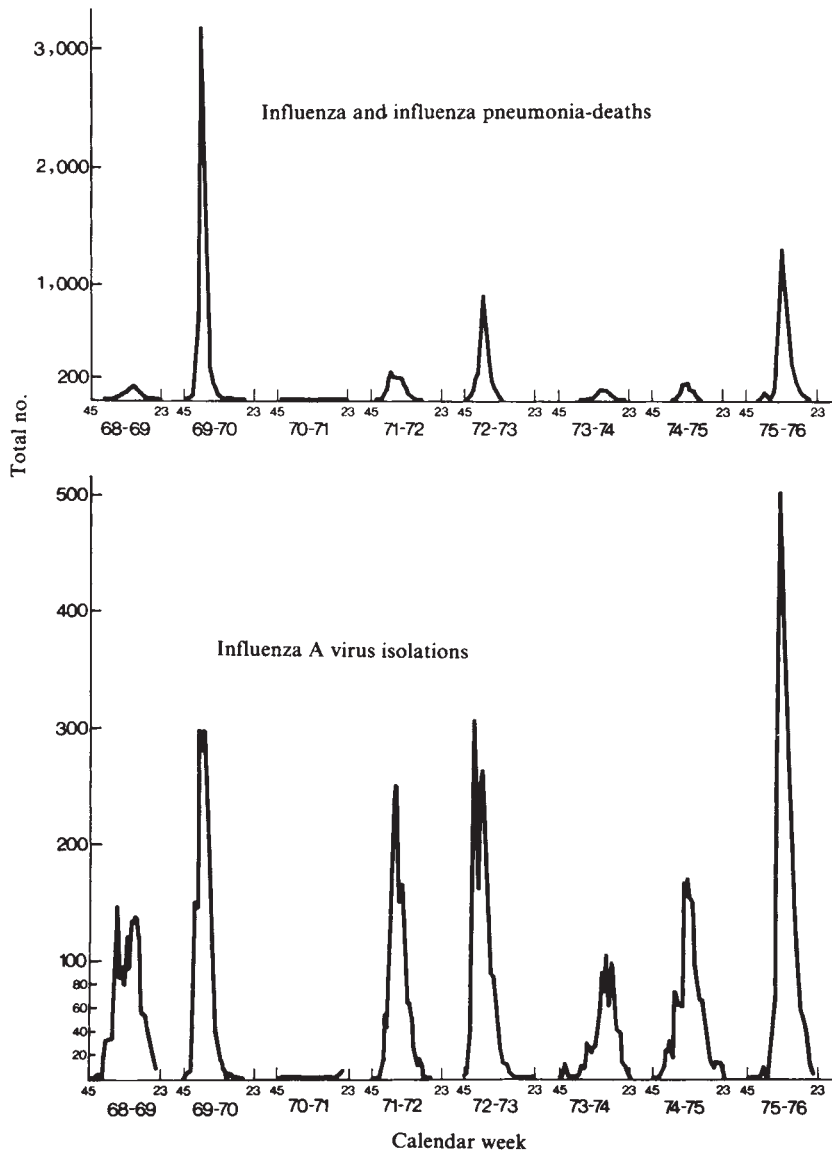


Fig. 1 The results of influenza surveillance in England and Wales for the years 1968–1976. The graphs were drawn from information provided by the Communicable Disease Surveillance Centre of the Public Health Laboratory Service, Colindale, London, part of which was published in detail in 1977 (ref. 29). Data of a similar nature for other countries up to 1973 are given in Assaad *et al.* (ref. 30).

Antigenic variation

Since the first isolation of an influenza virus from humans in 1933 (ref. 17), two major pandemics have been recorded (in 1957 and 1968) and on both occasions, the viruses responsible were characterized by marked differences in the antigenic properties of their haemagglutinins. These viruses were designated prototype strains of distinct antigenic subtypes: H2 in the case of the Asian viruses prevalent between 1957 and 1968 and H3 for the Hong Kong viruses which appeared in 1968 and are still in circulation¹⁸.

In addition to the pandemics which were observed initially within each period, epidemics were recorded which were thought to be caused by viruses which had undergone 'antigenic drift'¹⁹ from the prototypes. We discuss the structural basis of this process as it occurred between 1968 and 1975.

Antigenic variation in the Hong Kong (H3) viruses. In 1968 a strain of influenza virus, A/Hong Kong/1/68, was isolated and found to be antigenically unrelated to the viruses of the Asian (H2) subtype which were prevalent during the previous decade. Worldwide epidemics of disease due to infection with this virus were reported in the winters of 1968/9 and 1969/70, during which time the results of international surveillance indicated little variation in the antigenic properties of the viruses isolated²⁰ (Fig. 1). In the winter of 1971/72, an antigenically distinct virus, A/England/42/72 (similar to A/Memphis/102/72 in the USA), was isolated—this caused widespread epidemics in 1972/73. Several antigenically related strains were reported in 1973/74 and in 1974/75, and there

were further worldwide epidemics in 1975/76 following the appearance of viruses antigenically similar to A/Victoria/3/75 (ref. 21). This type of information was available from several countries—for the UK (data shown in Fig. 1), the number of deaths due to influenza and influenza pneumonia increased in 1969, 1972 and 1975, coincident with increases in the number of virus isolations and with the appearance of antigenically distinct viruses.

The amino acid substitutions in the 1972 and 1975 disease-causing strains were located in the three-dimensional structure of the 1968 haemagglutinin (Fig. 2). The amino acid sequences of the haemagglutinins of representative viruses isolated in 1968, 1972 and 1975 were determined from a combination of protein and nucleic acid sequence analyses^{4-6,10,13}. Fourteen amino acid substitutions were detected in the HA₁ polypeptide component of A/Memphis/102/72 relative to that of the A/Aichi/2/68 virus and a further 16 in HA₁ of A/Victoria/3/75. A summary of the structural locations and probable stereochemical consequences of these substitutions together with information from seven other partial sequences⁹ and analyses of the haemagglutinins of laboratory-selected antigenic variants^{15,16,22} is shown in Table 1.

Because the haemagglutinins of antigenic variants isolated in nature may also contain neutral substitutions, that is, mutations which do not affect the antigenic recognition by components of the immune system, two selection techniques have been used in attempts to identify those amino acid substitutions which influence antibody-antigen interaction. The work of Fazekas de

St. Groth and colleagues^{15,16,23} has involved the selection of variants by growth of an H3 1968 virus (A/NT/60/68) and its derivatives in the presence of the most avid fractions of hyperimmune antisera and, similarly, Laver, Webster and colleagues^{22,24,25} have selected variants of A/Memphis/1/71, a virus closely related to A/Hong Kong/1/68, by growth in the presence of monoclonal antibodies prepared against its haemagglutinin. After virus selection, amino acid sequence data were obtained for the haemagglutinins of the variants. Although only a partial sequence data are available, they indicate that variants selected using monoclonal antibodies contain single amino acid substitutions, which directly or indirectly alter an antigenic region of the haemagglutinin molecule sufficiently to prevent neutralization of the variant by the selecting antibody²². The locations of the amino acid substitutions in haemagglutinins of the laboratory-selected variants are summarized in Fig. 2c. **Suggested locations for the antigenic sites.** An antigenic site is a region of a molecule involved in antibody binding²⁶. Although there are no exhaustive immunochemical data such as those used in studies defining the antigenic sites on lysozyme and myoglobin^{27,28}, a combination of the amino acid sequence information and the three-dimensional structure of the 1968 haemagglutinin¹ allows identification of four regions of the Hong Kong haemagglutinin molecule in which amino acid substitutions seem to affect antibody binding.

Site A: an unusual protruding loop from amino acids 140 to 146, (see Fig. 2) which projects 8 Å from the local molecular surface, forms the centre of the most obvious antibody-binding site. Clearly the haemagglutinin of each antigenically distinct virus of epidemic significance has a mutation in this region (Fig. 2 and Table 1): A/Memphis/102/72: residue 144, Gly to Asp, and residue 122, Thr to Asn^{4,6}; A/Victoria/1/75 and A/Victoria/3/75: residue 145, Ser to Asn, and residue 137, Asn to Ser respectively^{9,12,13}. Furthermore, the haemagglutinins

of a group of monoclonal antibody-selected variants neutralization by single amino acid substitutions in this region: 143, Pro to His, Ser, Thr, Leu; 144, Gly to Asp, and 133, Asn to Lys²². Conserved amino acids Gly 134, Ala 138, Cys 139, Phe 147 and Phe (Tyr) 148 occurring at each end of the antigenically variable loop provide the structural foundation of the site in all 10 haemagglutinins examined (including one Avian³ and one Asian⁷ subtype).

Site B: this comprises the external residues 187–196 of an α helix, and adjacent residues along the upper edge of a pocket tentatively implicated in virus receptor binding¹ (see Fig. 2). Two substitutions in the haemagglutinin of the 1972 virus, A/Memphis/102/72: 188, Asn to Asp, and 155, Thr to Tyr^{4,6}; and two more in that of the 1975 virus, A/Victoria/3/75: 189, Gln to Lys and 193, Ser to Asn^{9,12,13} characterize the site, but the lack of sequence information for this region of the haemagglutinins of several variants leaves the comparison incomplete⁹. The substitution 186, Ser to Ile in the haemagglutinin of the 30D virus¹⁶, selected using 'avid-fraction' antisera, may also support the site designation. In addition, the existence of this structure in other haemagglutinins is suggested by the observed conservation of neighbouring residues. On the inward face of the helix residues 190 Glu, 191 Gln, 194 Leu and 195 Tyr are conserved in all the known sequences, as are 153 Trp and 154 Leu. However, because no amino acid substitutions in the haemagglutinins of variants selected by monoclonal antibodies comparable with those which support the delineation of site A have yet been reported, more direct evidence of antibody binding to this region is not available.

Site C: a bulge in the tertiary structure at the disulphide bond between Cys 52 and Cys 277, 60 Å from the distal tip of the molecule, comprises another antibody-binding site (see Fig. 2). The haemagglutinins of the viruses from both epidemic periods have substitutions clustered in this region:

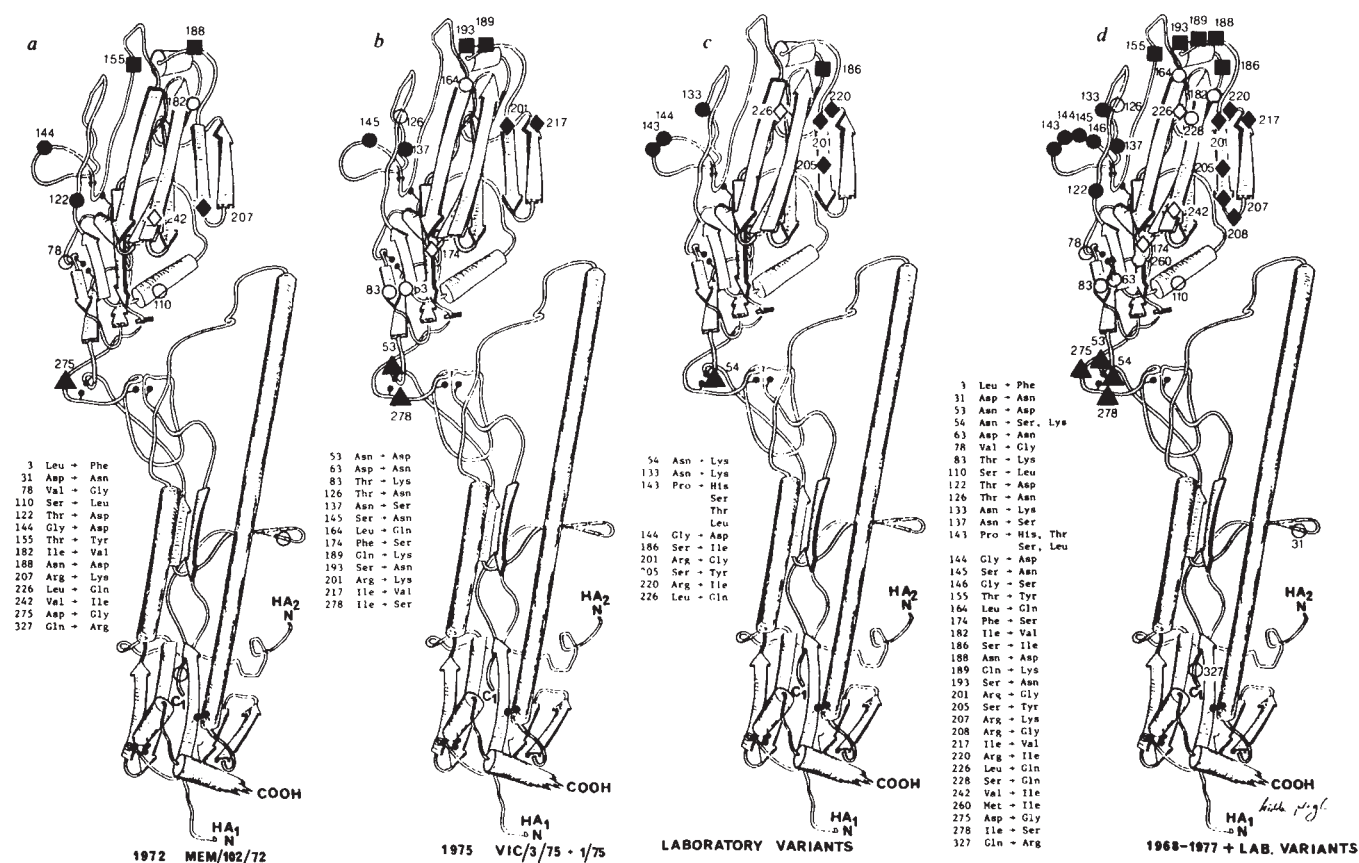


Fig. 2 Schematic drawing of a monomer of the AICHI-1968 haemagglutinin showing locations of amino acid substitutions on HA1. ●, Site A; ■, site B; ▲, site C; ◆, site D (subunit interface); ◇, surface; ○, neutral. *a*, Amino acid changes from AICHI/2/68 to MEM/102/72 (refs 4–6, 10, 11). *b*, Amino acid changes from MEM/102/72 to VIC/3/75 (refs 12, 13). *c*, Laboratory-selected antigenic variants from growth in monoclonal or most-avid-fraction antibody^{9,15,16,22}. (See Table 1 for details.) *d*, Summary of all known alterations from 1968 to 1977 including laboratory-selected variants^{3–16,22}. A three-dimensional (stereo) drawing of the same view of the molecule is found in Fig. 3 of ref. 1.

Table 1 Antigenic drift residues

Strain	Residue	Replacement (from)	Replacement (to)	Probable antigenic site	Structural location and probable effects of replacement
MEM/102/72*	144	Gly	Asp	A	External loop. Substitution is both larger and charged. Change from glycine may restrict ϕ , Ψ angles of main chain, altering the loop shape.
	122	Thr	Asn	A?	External extended chain (β). Substitution is larger, 18 Å from residue 144, and may effect the edge of a contact region centred at the 144 loop.
	155	Thr	Tyr	B	4 Å from residue 193 (VIC) and the C-terminal end of the 187–196 α -helix. Substitution is considerably larger and would project over the 'pocket'.
	188	Asn	Asp	B	N-terminal exterior face of α helix 187–196, 18 Å from residue 155 altered charge.
	275	Asp	Gly	C	Near 52–77 disulphide bond bulge. Asp 275 was H-bonded to Asn 53, an interaction that would be lost by the Gly substitution.
	207	Arg	Lys	D?	Subunit interface near residues 99 and 100 of another subunit. Substitution is conservative but shorter and unbranched.
	242	Val	Ile	D??	At the surface but near residues 205 and 207 of same subunit and 221 and 222 of an adjacent subunit, also near the oligosaccharide chain attached at 165. Substitution is larger.
	226	Leu	Gln	?	External on the front of the 'pocket', 8 Å from residue 137 (A), 8 Å from 186 (B?), 16 Å from 242 (D?). Substitution increases size and changes polarity. Gln could reach Ser 136 or Asn 38 for H-bonds. Found in 29C 'avid-fraction' variant. Reverts in VIC/3/75.
	3	Leu	Phe	None	Near membrane, probably inaccessible to antibodies. Reverts in VIC/3/75.
	327	Gln	Arg	None	100 Å from distal tip. Reverts in VIC/3/75.
	31	Asp	Asn	None	82 Å from distal tip.
	78	Val	Gly	None (E?)	Probably buried by oligosaccharide at 81.
	182	Ile	Val	None	Buried residue near 220.
	110	Ser	Leu	None	Near Arg 261, buried. Reverts in VIC/3/75.
	VIC/3/75† (VIC/1/75)‡	(145)	Ser	Asn	A
137		Asn	Ser	A?	External 7 Å from residue 145. Substitution is smaller, possibly near an edge of the A site.
189		Gln	Lys	B	External face of 187–196 α helix. Substitution is larger and charged.
193		Ser	Asn	B	External face of 187–196 α helix, 4 Å from 189 and 4 Å from 155. Substitution is larger with more H-bond potential. Possible new interaction with Tyr 155.
278		Ile	Ser	C	Surface residue in 52–277 disulphide bulge. Substitution is smaller and polar and could H-bond to Asn 54.
(53)		Asn	Asp	C	External in 52–277 disulphide bulge, next to 275. Introduces charge (VIC/1/75).
201		Arg	Lys	D?	Packed in the subunit interface against 217 of the other subunit, below 188. Substitution is smaller and less bulky and could effect the fit between subunits or the site 'B' helix or be recognized directly if the subunit interface were exposed by a conformational change.
217		Ile	Val	D?	Packed in the subunit interface against 201 of the other subunit. Similar to the 201 change. Under 189 (B).
174		Phe	Ser	D??	External after surface chain, 169 to 173: Pro, Asn, Asn, Asp, Asn. 14 Å from 242. 169 is next to 242.
164		Leu	Gln	None	Buried next to 127. Substitution gets larger and polar.
126		Thr	Asn	None	New oligosaccharide attachment site. Carbohydrate would cover a smooth face. Remote possibility that it could reach 133 or 122 and affect site A.
327				}	Same as 1968 AICHI strain.
3					
226					
110					
TEX/1/77§	63	Asp	Asn	None (E?)	New oligosaccharide attachment site, very near old site at 81.
	83	Thr	Lys	None (E?)	Destroys oligosaccharide attachment site at 81.
	146	Gly	Ser	A	External loop. Substitution increases size and potential for H-bonding to 135, 136 region, also limits Φ , Ψ angles of main chain (see MEM/102/72 residue 144).
	54	Asn	Ser	C	External as 53 above, next to 278 which is not yet sequenced in TEX/1/77.
ENG/187/70§	260	Met	Ile	None	Buried next to 177, near 174. Close to site D?
ENG/42/72§	228	Ser	Gln	None	In pocket near 226. Substitution is much larger, and does not reappear after this strain.
Monoclonally selected variant	208	Arg	Gly	D?	Near subunit interface, residue 207. Substitution smaller and lacks charge.
	143	Pro	His, Ser Thr, Leu	A	Loop. All these residue 143 variations were selected with a single antibody. Selection with other antibodies has produced identical sequence changes.
	144	Gly	Asp	A	Loop. (See MEM/102/72 residue 144).
	133	Asn	Lys	A	Selected with the same antibody that 143 Pro-Ser was selected.
	54	Asn	Lys	C	External 52–277 disulphide bulge. Substitution is larger and charged.
	205	Ser	Tyr	D	Buried in the subunit interface near 221 and 220 (See VIC/3/75 residue 201).
	Avid-fraction variant¶	30D	186	Ser	Ile
30D		220	Arg	Ile	D
34D	201	Arg	Gly	D?	Refer to VIC/3/65 residue 201.
	29C	226	Leu	Gln	?

* Refs 4, 5, 6.

† Refs 8, 9.

‡ Refs 12, 13.

§ Ref. 9.

|| Refs 23, 24.

¶ Refs 19, 20.

A/Memphis/102/72: 275, Asp to Gly^{4,6}; A/Victoria/1/75 and A/Victoria/3/75: 53, Asn to Asp and 278, Ile to Ser respectively^{9,12,13}. The absence of amino acid sequences for the 275 to 278 region of HA₁ from a number of partial sequences may also mean that this list is incomplete⁹. A substitution at position 54, Asn to Lys, in a variant selected by monoclonal antibody

prevents its neutralization²² and variant 30D also fails to react with the same monoclonal antibody¹⁶. Again the variable amino acids seem to be anchored in a matrix of conserved residues: 50 Lys, 51 Ile (Leu), 52 Cys, 277 Cys, 281 Cys and 286 Gly. The variable residues at all four positions are in direct contact. In the structure of the 1968 haemagglutinin, Asp 275 is hydrogen-

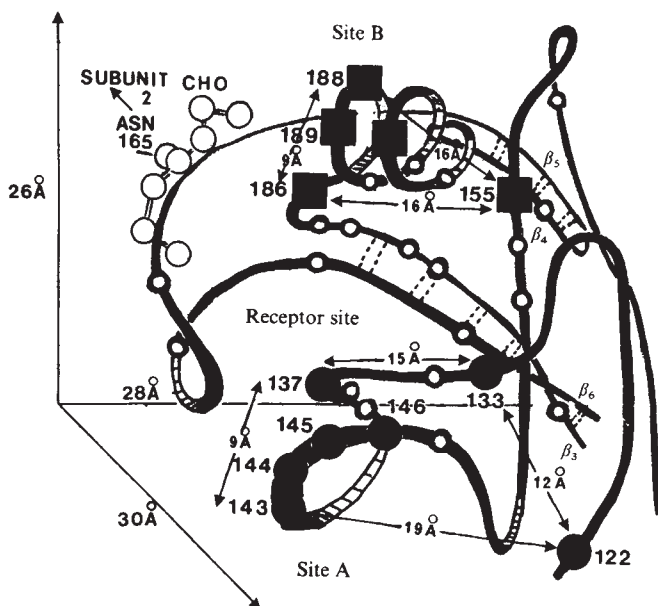


Fig. 3 Schematic drawing showing the size and shape of two of the proposed antigenic sites: ●, A and ■, B (see Fig. 2 and text). Their relationship to the pocket of conserved residues tentatively identified as the host sialic-acid-containing receptor binding site is shown.

bonded to Asn 53, Ile 278 is packed against Asn 54, and Cys 52 bonds to Cys 277. Changes in the types of contact between these residues (for details see Table 1) may be the structural alterations to which antibody binding is sensitive.

Site D: in the first three sites, the amino acid substitutions suggested as causing antigenic variation are external and, in principle, directly recognizable by molecules of the immune system. Site D departs from this apparently simple situation (site C may also). Several amino acid substitutions in the haemagglutinins of both natural and laboratory-selected antigenic variants occur in the interface region between subunits in the haemagglutinin trimer: A/Memphis/102/72: 207, Arg to Lys^{4,6}; A/Victoria/3/75: 201, Arg to Lys and 217, Ile to Val^{9,12,13}; monoclonal variant: 205, Ser to Tyr²², and 'avid-fraction' variants 34C: 201, Arg to Gly, and 30D: 220, Arg to Ile^{15,16} (see Figs 2, 4). These amino acids may be recognized directly as a result of a relative movement of the globular domains of HA₁ to expose the interface regions. However, it is also possible that the actual antibody binding site is remote from these amino acids but affected by the exact fit at the interface, which might be disturbed by the substitutions listed. A site which spans the monomer-monomer boundary or one near the oligosaccharide chain attached at residue 165, which overlaps that

boundary, could be particularly sensitive to alterations in the subunit contact region. Residues 242, Val to Ile and 226, Leu to Gln in A/Memphis/102/72^{4,5}; 174, Phe to Ser in A/Victoria/3/75^{9,12,13,25} and 226, Leu to Gln (also near sites A and B) in laboratory variant 29C (refs 15, 16) are likely components of this area (see Fig. 2). The amino acid sequences of peptides in the 174 and 242 regions are absent from many partial sequences⁹.

Other amino acid substitutions in HA₁ and HA₂ found in nature are considered to be antigenically neutral for reasons detailed in Table 1. For example, the substitution at residue 78, Val to Gly, in the A/Memphis/102/72 haemagglutinin, is buried by the oligosaccharide at position 81. Similarly, in A/Victoria/3/75 at residue 83, the Thr to Lys substitution destroys the oligosaccharide attachment site at residue 81, but the Asp to Asn change at 63 creates a new site—this means that the carbohydrate could occupy almost the same position.

The characterization of new variants may provide evidence for additional sites (for example, 122 and 126, site E?) or extend the size of those proposed.

Antibody interactions and neutralization. This analysis has not defined the extent of each site but only amino acid positions within sites that seem sensitive to substitution when examined serologically. Nevertheless, areas of the surface remain, not in the interfaces nor covered by oligosaccharide chains, that seem serologically unimportant. It is not clear why these regions are not immunogenic, although, for example, the concave shape of the pocket of conserved amino acids tentatively identified with receptor binding¹ probably prevents the production of avid antibodies against that region. The proximity of sites A, B and D to this tentative receptor site (Fig. 3) may be important in the mechanism of neutralization.

Table 1 summarizes changes in the size and/or change of the observed substitutions that would result in a decrease in the binding constant of antibody at a site as a result of steric repulsion or unfavourable charge interactions (the burying of an unpaired charge in the antibody-antigen complex). In addition to these mechanisms, the glycine-rich loop in site A protrudes in such a way as to make so few contacts with the rest of the structure that the effect of mutations in it may be amplified by altering the conformation of the entire loop. The more restricted range of torsion angles possible for amino acids replacing glycine and the potential for hydrogen and ionic bonds from the loop to adjacent regions of the molecule possible with polar or charged substitutions (Table 1) are candidates for such an alteration. (X-ray data on isomorphous crystals of singly substituted, monoclonally selected variant haemagglutinins prepared in collaboration with R. Webster are being collected to allow a direct visualization of these altered sites.)

Antigenic variation and disease incidence (1968–1975)

It seems from the above discussion that the haemagglutinins of both major disease-causing viruses of this period had at least one

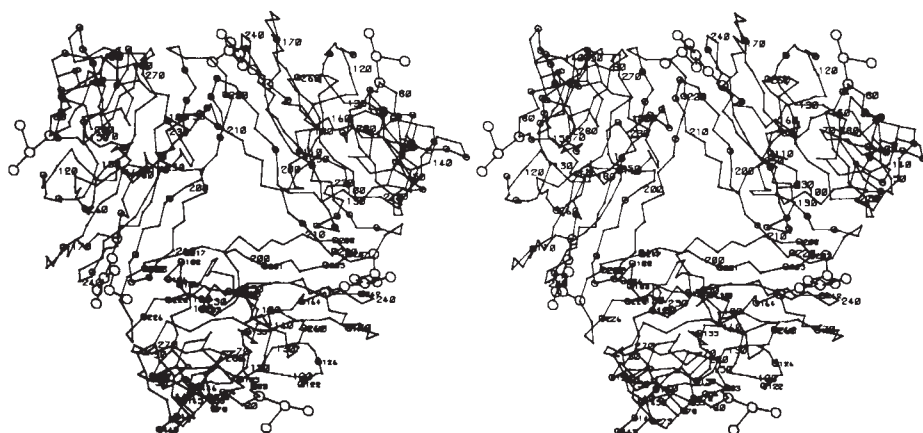


Fig. 4 Stereo drawing of the α -carbon tracing of residues 48–280 (HA₁) of the haemagglutinin trimer. Circled residues are natural and laboratory variations^{3–16,22} (see Fig. 2d). Tentative (see ref. 1) carbohydrate positions are shown as large open circles for the attachment sites at Asn 81 and Asn 165. The view is down the trimer axis, viewed from the outside towards the membrane.

mutation in each of the four antigenic sites described, and may have required these to avoid virus neutralization. Thus, the viruses acquired the antigenic potential to cause a recurrence of disease in individuals previously infected with viruses of the H3 subtype. Because other characteristics of viruses such as transmissibility and virulence are important in the timing and impact of epidemics, antigenic potential (the recreation of a threshold of susceptible hosts) may, in general, only be a necessary condition for, but not capable of, the generation of epidemics.

The possibility has also been raised²⁵ that one locus may be more immunogenic than others and therefore, when altered, might be of greater advantage to the virus. Site A, due to its accessibility, seems to be ideally suited both for avid antibody binding and for accepting amino acid substitutions without affecting the main framework of the molecule. To establish the relative importance of the four antigenic sites, a definition will be needed of the spectrum of antibody molecules in human sera with regard to amount, avidity and site specificity.

The character and location of amino acids on the 1968 haemagglutinin structure conserved in three subtypes³⁻¹⁴ indicates that the basic framework of the molecule is preserved¹. Because strictly conserved amino acids are found flanking each

of the proposed antigenic sites, these structures are also probably present in haemagglutinins from other subtypes. Whether or not the same sites are involved in the interaction of these other haemagglutinins with antibodies is unknown. In the haemagglutinin of other subtypes, these sites have not only varied considerably in amino acid sequence but probable points of oligosaccharide chain attachment are sometimes found within or near regions of the molecule sequentially equivalent to those proposed here as antigenic sites in the Hong Kong haemagglutinins: for example, in fowl plague virus³ (Hav1) at 133 (site A), 158 (site B) and 240 (possibly site D), and in A/Japan/305/57 (H2) (refs 7, 14) at 169 (site D). It is therefore possible that the immune recognition of such sites in the haemagglutinins of other subtypes is sterically prevented by oligosaccharide chains.

We thank Colin Ward and Theo Dopheide for A/Memphis/102/72 and X:31 A/Aichi/2/68 sequences, Walter Fiers for the sequence of A/Victoria/3/75 and X:31, and Mary Jane Gething and Mike Waterfield for Japan/305/57. We thank Andrew Chersonson for assistance in sequence searches and comparisons, and acknowledge support from NIH grant AI 13654 (D.C.W.). Computing was supported by NSF PC-77-11398.

Received 20 October; accepted 18 November 1980.

1. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366-373 (1981).
2. Webster, R. G. & Laver, W. G. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) 269-314 (Academic, New York, 1975).
3. Porter, A. G. *et al. Nature* **282**, 471-477 (1979).
4. Ward, C. W. & Dopheide, T. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 27-38 (Elsevier, Amsterdam, 1980).
5. Dopheide, T. A. & Ward, C. W. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 21-26 (Elsevier, Amsterdam, 1980).
6. Sleight, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 69-78 (Elsevier, Amsterdam, 1980).
7. Gething, M. J., Bye, J., Skehel, J. & Waterfield, M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 1-10 (Elsevier, Amsterdam, 1980).
8. Min Jou, W. *et al. in Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 63-68 (Elsevier, Amsterdam, 1980).
9. Verhoeyen, M. *et al. Nature* **286**, 771-776 (1980).
10. Ward, C. W. & Dopheide, T. A. *Biochem. J.* (submitted).
11. Threlfall, G., Barber, C., Carey, N. & Emtage, S. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 51-61 (Elsevier, Amsterdam, 1980).
12. Min Jou, W. *et al. Cell* **19**, 683-696 (1980).
13. Gething, M. J., Bye, J., Skehel, J. & Waterfield, M. *Nature* **287**, 301-306 (1980).
14. Both, G. W., Sleight, M. J., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 81-89 (Elsevier, Amsterdam, 1980).
15. Moss, B. A., Underwood, P. A., Bender, V. J. & Whittaker, R. G. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 329-338 (Elsevier, Amsterdam, 1980).
16. Smith, W., Andrewes, G. H. & Laidlaw, P. P. *Lancet* **ii**, 66-68 (1933).
17. World Health Organization (1971) *Bull. Wld Hlth Org.* **45**, 119-124 (1971).
18. Burnett, F. M. *Principles of Animal Virology*, 382 (1955).
19. Pereira, M. S. *Br. med. Bull.* **35**, 9-14 (1979).
20. Stewart-Harris, C. *Br. med. Bull.* **35**, 3-8 (1979).
21. Laver, W. G. *et al. in Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 295-306 (Elsevier, Amsterdam, 1980).
22. Fazekas de St. Groth, S. *Topics Infect. Dis.* **3**, 25-48 (1978).
23. Laver, W. G. *et al. Virology* **98**, 226-237 (1979).
24. Webster, R. G. & Laver, W. G. *Virology* **104**, 139-148 (1980).
25. Atassi, M. Z. & Smith, J. A. *Immunochemistry* **15**, 609-610 (1978).
26. Atassi, M. Z. *Immunochemistry* **15**, 909-936 (1978).
27. Atassi, M. Z. *Immunochemistry* **12**, 423-438 (1975).
28. Public Health Laboratory Service Standing Advisory Committee on Influenza *J. Hyg., Camb.* **78**, 223-233 (1977).
29. Assad, F. A., Cockburn, W. C. & Sundaresan, T. K. *Bull. Wld Hlth Org.* **49**, 219-233 (1973).

A gene chimaera of SV40 and mouse β -globin is transcribed and properly spliced

Gilbert Chu & Phillip A. Sharp

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

A gene chimaera of the first exon from simian virus 40 (SV40) sequences coding for T antigen placed upstream from the third exon from mouse β -globin, and separated by the resultant new chimaeric intron, was cloned into a bacterial plasmid. When transfected into monkey cells, the gene chimaera was transcribed, polyadenylated and spliced using the donor splice site from SV40 and the acceptor splice site from mouse β -globin. This result suggests that a donor site from one gene can be spliced to an acceptor site from another gene.

THE finding of intervening sequences within mammalian genes has led to speculation that they take part in both phylogeny and ontogeny by facilitating the rearrangement of genes. These ideas are based on the assumption that the cell can properly splice together RNA from any two exons. Thus, evolution and differentiation may proceed in quantum jumps as a new gene is built by rearrangement of exons coding for functional protein domains from different pre-existing genes¹. Differentiation of the immunoglobulin gene proceeds in such a manner². Similarly, the regulation of a gene may be radically altered if a promoter

sequence and short exon are introduced upstream from a series of exons coding for a complete protein. Insertion of this transposable promoter could occur thousands of nucleotides from the protein-coding exons.

If exons can be shuffled arbitrarily and the RNA is properly spliced, then there should be similarity among all donor sites and among all acceptor sites. In fact, the boundaries of intervening sequences are characterized by two consensus sequences, one for donor sites and one for acceptor sites³. Furthermore, there are examples of splicing pattern in which two or more donor sites

are spliced to a common acceptor site (for example, the early region of SV40)^{4,5}. The sequences of such donor sites seem to bear no special relationship other than their relationship to the consensus donor sequence. There are similar examples of splicing patterns in which two or more acceptor sites are spliced to a common donor (for example, the late mRNAs of adenovirus)^{6,7}. These observations support, but do not prove, the hypothesis that a donor site from one gene can be spliced to an acceptor site from another gene. The experiment described here tests this hypothesis.

SV40-mouse β -globin chimera

A gene chimera of the first exon from SV40 T antigen placed upstream from the third exon from mouse β -globin (M β G), and separated by the resultant new chimaeric intron, was cloned into the plasmid vector pBR322 (see Figs 1, 2). The recombinant genome DNA was isolated from bacteria and transfected into African green monkey kidney cells (CV-1 cell line). RNA was collected from the cells and analysed by a modification of the S₁-endonuclease gel technique⁸—hybridization in 80% formamide to an end-labelled DNA probe, followed by digestion with S₁ endonuclease and electrophoresis on either neutral or denaturing gels.

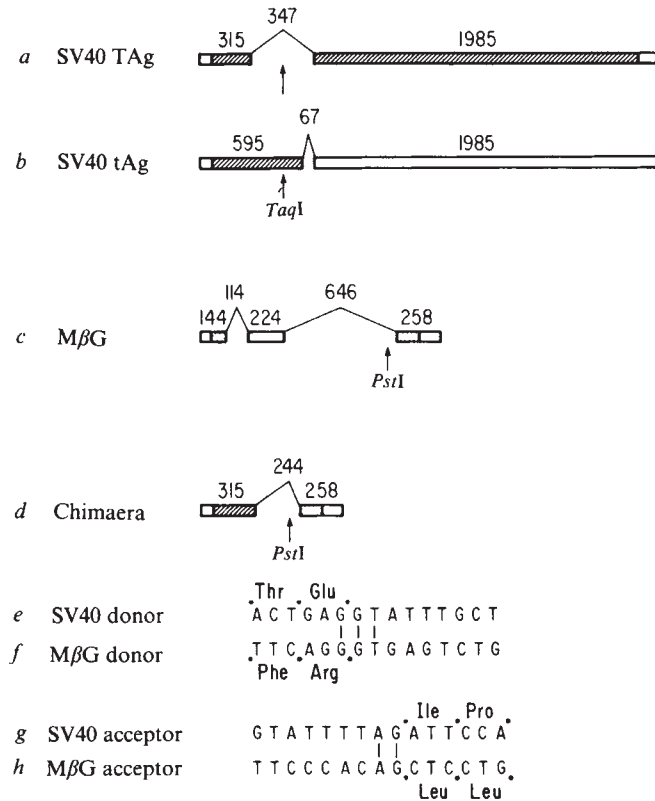


Fig. 1 mRNAs from early SV40, M β G and the gene chimera. *a, b*, The two major SV40 mRNAs produced early in the lytic cycle, corresponding to large T and small t antigens. The coding regions are hatched. *c*, The mouse β -globin mRNA where the coding regions are dotted¹⁵. *d*, The chimera mRNA which would be transcribed from a gene containing SV40 sequences upstream from the SV40 *TaqI* site and containing M β G sequences downstream from the M β G *PstI* site. The ligation is achieved by introducing *PstI* linkers (Collaborative Research) to the *TaqI* site in SV40 and annealing to the *PstI* site in M β G. The chimeraic mRNA contains the first SV40 T-antigen exon and the third M β G exon, with the two exons separated by a chimaeric intron containing both SV40 and M β G sequences. *e-h*, The splice junctions from SV40 and M β G. The junction sequences from the SV40 T-antigen intron *e* and *g*, and the second M β G intron *f* and *h*¹⁵ are shown. The chimeraic intron consists of SV40 donor and M β G acceptor. Note that the phase of the reading frame relative to the splice junction is the same for the SV40 and M β G introns. Thus the spliced chimeraic mRNA preserves the proper reading frame and permits the translation of a chimeraic protein.

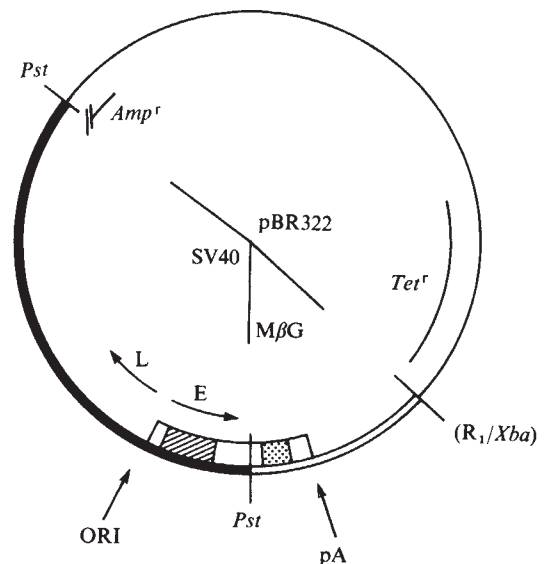


Fig. 2 Structure of the SV40-M β G-pBR322 recombinant genome. SV40 sequences (2,460 base pairs) are indicated by the solid bar, mouse β -globin sequences (~900 base pairs) by the single line and pBR322 plasmid sequences (3,607 base pairs) by the double line. The mouse β -globin (M β G) gene has been sequenced¹⁵. The SV40 DNA was made from strain 777. Strain 776 has been sequenced^{21,22}. Strains 776 and 777 show no difference, by S₁ digestion, in the SV40 early region of interest in this work (G.C., unpublished data). The M β G-pBR322 junction was formed by the blunt-end ligation²³ of the *XbaI* site in the 3'-flanking sequence of M β G with the *EcoRI* site in pBR322. The SV40-M β G junction was formed by cleavage of SV40 at the *TaqI* site, filling in the staggered end to form a blunt end by Klenow fragment of DNA polymerase I, attachment of *PstI* linkers to the blunt end, digestion by *PstI* and staggered-end ligation to the *PstI* site in the second intron of M β G. The SV40-pBR322 junction was formed by the staggered-end ligation of the *PstI* site at 0.04 map units in SV40 to the *PstI* site in the ampicillin-resistance gene of pBR322. The structure was confirmed by analysis with restriction enzymes *HindIII* and *BglII*, and by double digests with *HindIII* and *PstI*, *BamI* and *BglII*, and *HindIII* and *BamI*. The restriction patterns confirmed expectations from the construction. In particular, the region of the chimeraic intron was shown to contain a *PstI* site and to be of the length predicted from the SV40 and M β G sequences, to an accuracy of 20 base pairs. The recombinant plasmid was grown in strain C600 of *Escherichia coli* under P2 containment. In SV40, the origin of replication (ORI) and the directions of early (E) and late (L) transcription are indicated. In M β G, the site of polyadenylation (pA) is shown¹⁵. The DNA sequences coding for the chimeraic mRNA are indicated by the broad bar. The hatched area indicates the coding region of the first exon in SV40 T antigen. The dotted area indicates the coding region of the third exon in M β G. The two coding regions are separated by a chimeraic intron. In pBR322, the ampicillin (*Amp*^r) and tetracycline (*Tet*^r) resistance genes are indicated by the solid lines. The *Amp*^r gene is no longer intact.

Chimaeric mRNA is transcribed, polyadenylated and spliced using the normal M β G acceptor site

The chimeraic genome was transfected into CV-1 cells by a modification of the calcium phosphate co-precipitate technique⁹. It was found that 15% of the cells could be routinely transfected with SV40 DNA in this manner (G.C., unpublished data). To obtain optimal levels of chimeraic mRNA, the chimeraic genome was co-transfected with DNA consisting of SV40 cloned into pBR322 at the *BamHI* site. The A gene product is necessary for the replication of SV40 (ref. 10), and a 10-fold increase in early mRNA is observed between 9 h and 48 h after infection with SV40 virus^{11,12}. The SV40-pBR322 recombinant contains an intact early region, and thus replication from the SV40 origin of replication will take place. As the chimeraic genome also contains the SV40 origin of replication, it was hoped that co-transfection would produce a similar increase in chimeraic mRNA, and such an increase was observed. RNA was collected 48 h after transfection. In some cases, poly(A)⁺ RNA was selected by passage over an oligo(dT) column; all poly(A)⁻ fractions were treated with DNase I. A DNA probe

labelled at the 5' end (see Fig. 3) was hybridized to the RNA preparations in 80% formamide at 47 °C, treated with S_1 endonuclease and the products electrophoresed.

The results of hybridization of the RNA with the DNA probe labelled at the 5' end of the *BalI* site are shown in Fig. 4 for a neutral 2% agarose gel (a) and for a denaturing 7 M urea 8% acrylamide gel (b). *BalI* cleaves in the M β G sequences in the chimaera and thus a DNA probe labelled at this site will only hybridize to RNA complementary to globin sequences and not to RNA transcribed from the complementing SV40. The migration of a band on a neutral gel indicates the distance of the 5' end of the mature RNA from the 5'-end label of the probe. RNA transcribed from the chimaeric genome that was spliced would generate a 415-base pair RNA/DNA hybrid (see Fig. 3). Some 315 and 100 base pairs of this hybrid are contributed by the exons from SV40 and M β G, respectively. Bands, generated by the poly(A)⁺ fraction from cells transfected with the chimaeric genome after hybridization to DNA labelled at the *BalI*

probe to globin mRNA (Fig. 4b, lane 9). A band of the same length, 100 nucleotides, was produced by hybridization to poly(A)⁺ RNA extracted from a cell transfected with the chimaera. Because a sequencing ladder was run in parallel lanes and because hybridization of the probe to both chimaeric mRNA and M β G RNA produced co-migrating bands, we can assume that the acceptor sites for the chimaeric and M β G splices are identical, 100 nucleotides upstream from the *BalI* site in the third M β G exon (see Fig. 3).

Chimaeric mRNA is polyadenylated at the normal M β G polyadenylation site and spliced using the normal SV40 donor site

Hybridization results with the probe (labelled at the 3' end of the *HindIII* site in SV40 sequences) are shown in Fig. 5 for a neutral 2% agarose gel (a) and for a denaturing 7 M urea 8% acrylamide gel (b). The migration of a band on a neutral gel indicates the distance of the 3' end of the mature mRNA from the 3'-end label of the DNA probe. Hybridization of a probe labelled at the 3' terminus of the *HindIII* site to a spliced chimaeric RNA would generate a 511-base pair RNA/DNA hybrid (see Fig. 3). As shown in Fig. 5a, lane 4, a band migrating as a 520 $\pm$ 20-base pair RNA/DNA hybrid resulted from hybridization to poly(A)⁺ RNA extracted from cells transfected with the SV40 M β G recombinant.

Because the RNA in this experiment was collected from cells co-transfected with both cloned chimaeric DNA and cloned SV40 DNA, and because the probe is labelled at the 3' end in the SV40 sequence, there are bands in Fig. 5a, lane 4, corresponding to the early SV40 mRNAs for small t and large T antigens. This can be seen by comparing lane 4 with lane 13, which is the result of hybridization to RNA from cells infected with SV40 virus alone. The bands in Fig. 5a, lane 4 for small t and large T antigen are much lighter than the band for the chimaeric mRNA because the T_m for the region of SV40 between the origin of replication and *TaqI* sites is ~ 36 °C (ref. 5) and the T_m for the chimaeric DNA segment is 46 °C. The hybridization was done at 47 °C and therefore strongly favoured the detection of chimaeric mRNA over SV40 early mRNAs.

The migration of a band on a denaturing gel indicates the distance of the first donor splice site from the 3'-end label. The *HindIII* site of SV40 is some 256 nucleotides from the donor site in the mRNA for the large T antigen. The denaturing gel in Fig. 5b shows that the position of the donor site for the chimaeric splice is at the usual position of the donor for SV40. Part of the same sample analysed on the neutral gel in Fig. 5a, lane 4, was loaded on to the denaturing gel in Fig. 5b, lane 4. Hybridization of the probe to both chimaeric (Fig. 5b, lane 4) and SV40 mRNA (Fig. 5b, lane 6) produced bands which co-migrated at a rate consistent with 254 $\pm$ 5 base pairs from the donor splice site to the 3' end of the probe. The markers in lane 1 include a doublet at 276 and 284 base pairs. Comparison of the co-migrating bands with the distance separation of the doublet shows that the donor sites for the chimaeric and SV40 splices are the same, up to an uncertainty of two nucleotides.

Note that Fig. 5b, lane 4, is the result of hybridization of the chimaeric probe to RNA from cells co-transfected with both cloned chimaeric and cloned SV40 DNA. When the heteroduplexes are run on a denaturing gel, the bands corresponding to chimaeric mRNA and SV40 large-T mRNA should co-migrate and there will be a second band corresponding to SV40 small-t mRNA. We conclude that the band in Fig. 5b, lane 4, is almost entirely due to chimaeric mRNA and only receives a small contribution from SV40 large-T mRNA for two reasons. First, the result in Fig. 5a, lane 4, shows that the level of hybridization to chimaeric mRNA is much higher than that for mRNA for either SV40 large T and small-t antigens. Second, the small t and large T mRNAs are present in roughly equal amounts in Fig. 5a, lane 4, and the small-t mRNA is represented by a very faint band in Fig. 5b, lane 4.

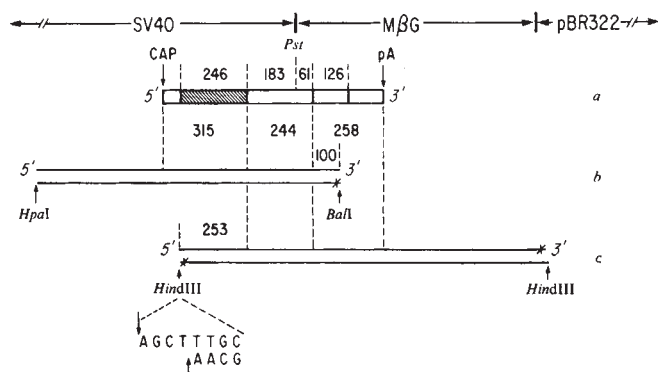


Fig. 3 Gene chimaera and the DNA probes used to analyse the chimaeric mRNA. *a*, The gene chimaera (817 base pairs) is shown from the major initiation site in SV40 (CAP) to the polyadenylation site in M β G (pA). The chimaeric intron is 244 base pairs long, containing 183 base pairs of SV40 sequence and 61 base pairs of M β G sequence. The two coding regions contain 372 base pairs, corresponding to a polypeptide of 124 amino acids. Note that both the coding region and the donor splice site of SV40 t antigen extend beyond the SV40-M β G junction. At the top of the figure the regions corresponding to SV40, M β G and pBR322 sequences are delineated. *b*, the 5'-end-labelled DNA probe was prepared by cutting the recombinant genome with *BalI*, labelling with [γ -³²P]ATP in the presence of t_4 polynucleotide kinase, cutting with *HpaI* and isolating the fragment by gel elution²⁴. *c*, The 3'-end-labelled DNA probe was prepared by cutting the recombinant genome with *HindIII*. The DNA was then incubated with dCTP and T₄ DNA polymerase in 10 μ l of a buffer of 50 mM Tris pH 8, 7 mM MgCl₂, 7 mM β -mercaptoethanol at 9 °C to allow the 3' $\rightarrow$ 5' exonuclease activity of the enzyme to remove nucleotides up to the first position containing a cytosine²⁵. After 30 min, [α -³²P]dATP was added to allow the 5' $\rightarrow$ 3' polymerase activity of the enzyme to add labelled nucleotides. After labelling, the fragment was isolated by elution from an agarose gel. The left-hand sequence after *HindIII* cleavage is indicated.

site, are shown in Fig. 4a, lane 4. An RNA/DNA hybrid migrating at a rate expected for a 420 $\pm$ 10 base pairs is clearly present. No such band appears in either the poly(A)⁻ fraction (Fig. 4, lane 5) or a mock transfection sample (Fig. 4, lane 3).

The same experiment was done with RNA collected 9 h after transfection with chimaeric DNA alone. About one-tenth the amount of chimaeric RNA was observed in these conditions as compared to co-transfection with SV40. On a neutral gel, the same size band of 420 $\pm$ 10 base pairs was observed in both the total cytoplasmic sample and the poly(A)⁺ fraction, but not in the poly(A)⁻ fraction or in a mock transfection cytoplasmic sample (data not shown).

The migration of a band on a denaturing gel indicates the distance of the first acceptor splice site from the 5'-end label. Hybridization of DNA labelled at the *BalI* site to RNA spliced at the acceptor sequences on the third M β G exon would yield a S_1 -resistant segment some 100 nucleotides in length. A band migrating at this rate was generated by hybridization of the

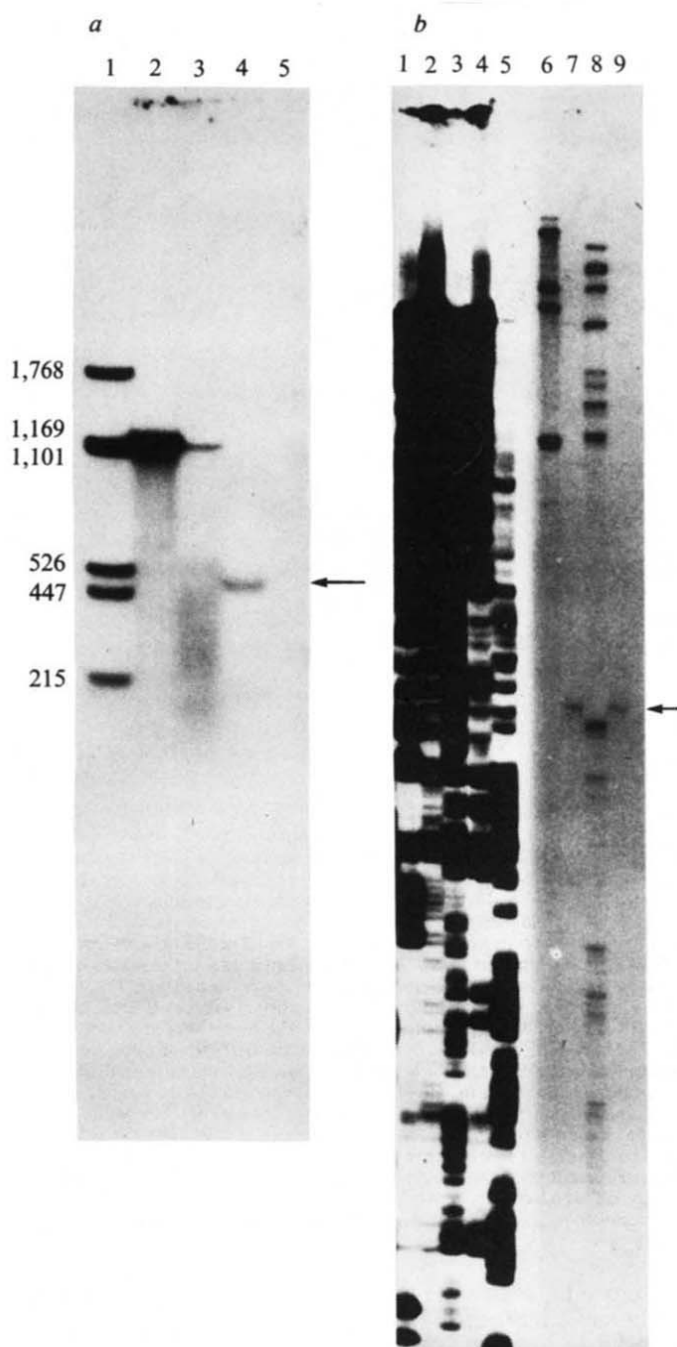


Fig. 4 Analysis with 5'-end-labelled probe of the RNA from cells co-transfected with the gene chimaera and cloned SV40 DNA. The probe is labelled at the *BalI* site of the chimaeric clone with polynucleotide kinase (see Fig. 3b) and hybridization was performed in 80% formamide at 47 °C (ref. 8). The transfection procedure is a modification of that used by Graham and van der Eb⁹ and Frost and Williams²⁶. Here, the cells are exposed to a calcium phosphate/DNA co-precipitate while in suspension (G.C., in preparation). *a*, Neutral 2% agarose gel. Lane 1 shows DNA markers, SV40 DNA cleaved by *HindIII* resulting in fragments of length 1,768, 1,169, 1,101, 526, 447 and 215 base pairs. Lane 2 shows intact probe. Lane 3 is the result of hybridization to the cytoplasmic RNA from mock-transfected cells. Lanes 4 and 5 are the poly(A)⁺ and poly(A)⁻ cytoplasmic fractions of cells co-transfected with the gene chimaera and cloned SV40 DNA, collected after 48 h. The band in lane 4 corresponds to a heteroduplex of 420 ± 20 base pairs. This value is determined by analysis with heteroduplex markers, as RNA-DNA heteroduplexes migrate slightly slower than DNA-DNA duplexes on a neutral gel (see data in Fig. 5). *b*, 7 M urea 8% acrylamide gel. Lanes 1-5 show a sequencing ladder used as markers. Lane 6 is a marker lane (SV40 cleaved with *HindIII*). Lane 7 is the result of hybridization to RNA from co-transfected cells. Lane 8 is a marker lane. Lane 9 is the result of hybridization to MβG mRNA. The bands in lanes 7 and 9 are co-migrating, with an uncertainty of less than one base pair on the sequencing ladder. Because the sequencing ladder in lanes 1-5 contained much more specific activity than the experimental data in lanes 6-9, the gel was photographed to give optimal exposure in lanes 6-9, while preventing overexposure in lanes 1-5.

Analysis of nuclear RNA from co-transfected cells reveals the existence of RNA which is unspliced and which migrates with a size consistent with that expected for the unspliced precursor to the chimaeric message. The RNA observed appears in both poly(A)⁺ and poly(A)⁻ fractions and is not observed in the cytoplasm (Fig. 5a, lane 11). In addition, two other species of RNA are observed, one which protects the full length of the 5'-end-labelled probe, terminates at the MβG polyadenylation site and is unspliced, and the other which migrates as if it were the spliced product of the first, with the proper chimaeric intron excised (data not shown). These RNA species represent transcripts which must initiate upstream of the *HpaI* site, either in SV40 or pBR322. As before, the spliced mRNA is found only in the poly(A)⁺ fraction, whereas the unspliced RNA is found in both poly(A)⁺ and poly(A)⁻ fractions. Interestingly, neither spliced nor unspliced species accumulate in the cytoplasm.

RNA splicing

We have shown that a gene chimaera of SV40 and MβG is transcribed and spliced with the proper use of the SV40 donor and the MβG acceptor. Because the exons used were obtained from such different sources, monkey virus and mouse, and because the sequences show so little homology between the two donors and between the two acceptors (see Fig. 1), we conclude that splicing can occur over a wide range of donors and acceptors. In the particular case studied here, the spliced mRNA conserves the reading frame (see Fig. 1), and it remains to be seen whether proper splicing will occur if a chimaera is constructed where the spliced mRNA does not conserve the reading frame.

Recent studies of expression of mutated transcription units have suggested the interchangeability among donor sites or among acceptor sites. The mouse κ -chain immunoglobulin gene consists of three exons. In an aberrant κ gene, the donor site at the second exon is deleted and the transcribed message is formed by joining the first and third exons. This message translates into a polypeptide as the two exons have the same reading phase relative to the site of splicing¹³. Similarly, some mRNAs expressed after infection of cells with an Ad2-SV40 hybrid virus, Ad2⁺ND4, seem to be the product of splicing adenovirus RNA sequences to SV40 sequences¹⁴. (The term 'chimaeric splice' was used in the title of this article.)

Interchangeability among donor sites and among acceptor sites raises several questions. The two donors from the first and second intervening sequences in MβG show considerable homology to each other¹⁵, and similarly the two acceptors. But the improper splice from the first exon to the third exon has not been observed^{16,17}. If in the chimaeric gene the MβG acceptor is spliced to the SV40 donor, which bears little resemblance to the usual MβG donor, why does the improper splice in MβG not occur?

The most attractive explanation is that splicing activity operates in a processive way. That is, it starts at one end of the intervening sequence, for example, at the acceptor site, and completes the splice when it reaches the first donor sequence. This restricts RNA splicing to an intramolecular process and insures against production of a nonfunctional protein by skipping an exon. A paradox immediately arises as there are transcription units with multiple donor sites for a single acceptor site, for example, the two SV40 early messages^{4,5}. Also, the adenovirus late mRNAs have multiple acceptor sites for a single donor site^{6,7}. A processive splicing mechanism is consistent with such splicing patterns only if splicing activity either (1) reaches the first possible splice termination site and then restarts the splice, or (2) always begins at the donor site in some transcription units and at the acceptor site in others.

The excision of a single intervening sequence may proceed in a processive manner, but it is clear from studies of ovalbumin¹⁸ and vitellogenin nuclear RNA¹⁹, that the excision of a series of intervening sequences does not necessarily proceed in a specific

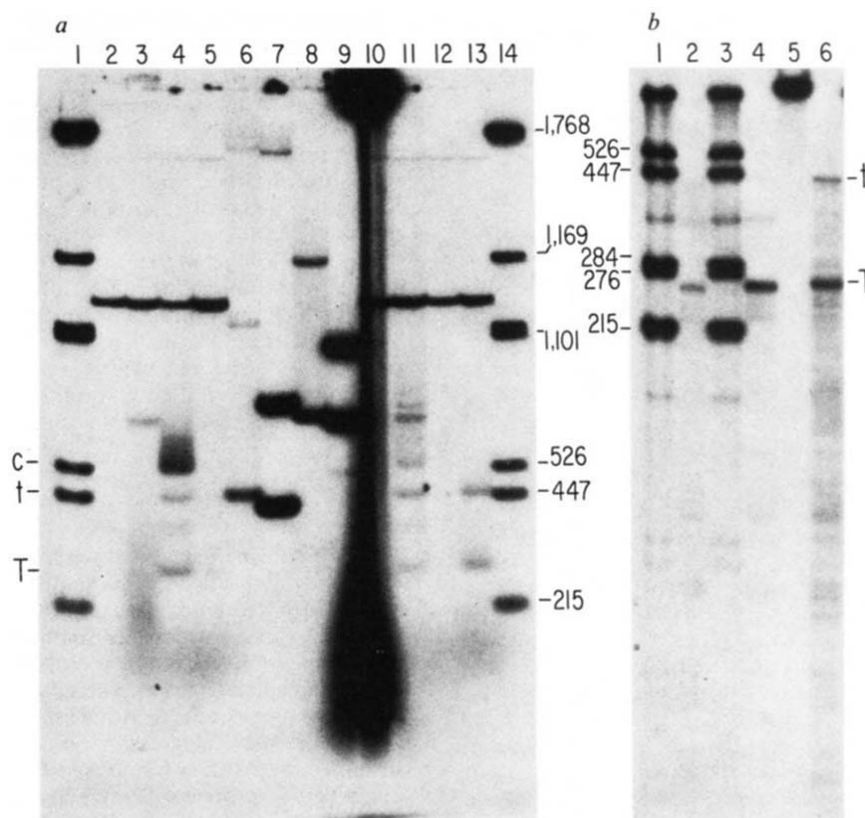


Fig. 5 Analysis with 3'-end-labelled probe of the RNA from cells co-transfected with the gene chimaera and cloned SV40 DNA. The probe is labelled at the *Hind*III site of the chimaeric clone with T_4 DNA polymerase (see Fig. 3c) and hybridization was done at 47 °C. *a*, Neutral 2% agarose gel. Lane 1 shows DNA markers, SV40 cut by *Hind*III. Lane 2 shows intact probe. Lane 3 is cytoplasmic RNA from mock-transfected cells. The band in lane 3 migrating at a rate corresponding to 710 base pairs is an artefact sometimes observed in this series of hybridizations. Lanes 4 and 5 are the poly(A)⁺ and poly(A)⁻ cytoplasmic fractions of cells co-transfected with the gene chimaera and cloned SV40 DNA, collected after 48 h. Marked by arrows to the left of the gel are the expected migration patterns of heteroduplexes formed with chimaeric mRNA (c), SV40 small t antigen mRNA (t) and SV40 large T antigen mRNA (T). Note the three bands in lane 4 which migrate at the predicted rates for c, t and T. Lane 10 is nuclear RNA from mock-transfected cells. Lanes 11 and 12 are the poly(A)⁺ and poly(A)⁻ nuclear fractions of co-transfected cells. There are bands present corresponding to the artefact in lane 3, and to c, t and T. There is also an extra band migrating at 750 ± 20 base pairs, consistent with the 755 base pairs predicted for an unspliced RNA terminating at the M ϕ G polyadenylation site. Lane 13 is RNA from cells infected with SV40 virus. There are bands present which correspond to mRNAs for t and T. Lane 14 shows the same markers used in lane 1. Lanes 6 and 7 are results with SV40 DNA labelled at the *Bcl*I site and subsequently cleaved with *Pst*I. Lane 7 shows intact DNA prepared in this way and lane 6 shows the result of hybridization of the labelled DNA to cRNA prepared from SV40 DNA with *E. coli* DNA polymerase. Note that the DNA-RNA heteroduplex in lane 6 migrates slightly slower than the DNA-DNA duplex in lane 7. Lanes 8 and 9 show a similar effect to that of lanes 6 and 7, but with SV40 DNA labelled at the *Bcl*I site and subsequently cleaved with *Hind*III. *b*, 7 M urea 8% acrylamide gel. Lanes 1 and 3 show DNA markers, SV40 cut by *Hind*III, end-labelled by T_4 polynucleotide kinase and recut by *Pst*I, resulting in fragments of length 1,402, 1,101, 911, 526, 447, 284, 276 and 215 base pairs. Lane 2 shows the poly(A)⁺ nuclear fraction of co-transfected cells from the same hybridization sample loaded on to the neutral gel in Fig. 5a, lane 11. Lane 4 shows the poly(A)⁺ cytoplasmic fraction of co-transfected cells from the same hybridization sample loaded on to the neutral gel in Fig. 5a, lane 4. Lane 5 shows the 3'-end-labelled DNA probe. Lane 6 shows the result of hybridization to RNA from cells infected with SV40 virus. (The probe used in lane 6 was the probe used for the hybridization in lanes 2 and 4, but cleaved with *Taq*I. Because the region of SV40 between the *Bgl*II and *Taq*I sites has a T_m of only 36 °C, the hybridization was done at 40 °C.) The bands in lanes 2 and 4 are co-migrating to an accuracy of less than two base pairs.

order. Thus the RNA splicing activity must be capable of interacting in a processive manner in the middle of a long RNA chain containing multiple intervening sequences.

A general model for splicing of pre-messenger RNA would consist of the recognition by a molecular structure of a set of RNA sequences, the consensus sequence, near either the donor or acceptor site, and then translocation of the structure along the intervening sequence until termination occurs at either the acceptor or donor site, respectively. The obvious similarity between this process and translation of mRNA by ribosomes is striking. Perhaps a nuclear analogue of the ribosome, that is, an entity containing both structural RNA and multiple proteins,

will prove to contain the RNA splicing activity. The studies of Steitz and colleagues on nuclear RNPs are pertinent in this regard²⁰.

Andrew Buchman and Paul Berg (personal communication) have constructed a chimaeric gene similar to that described here and have obtained similar results from the study of RNA transcribed from it.

We thank our colleagues for helpful discussions. P.A.S. was supported by American Cancer Society grant MV-37D, NSF grant PCM78-23230 and NIH Program Project grant CA-26717. This work was partially supported by Cancer Center Core grant CA-14051 from the NIH.

Received 26 August; accepted 12 November 1980.

- Gilbert, W. *Nature* **271**, 501 (1978).
- Hozumei, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
- Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4853-4857 (1978).
- Lebowitz, P. & Weissman, S. M. *Curr. Topics Microbiol. Immun.* **87**, 42-172 (1979).
- Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1274-1278 (1978).
- Berget, S. M., Moore, C. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3171-3175 (1977).
- Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. *Cell* **12**, 1-8 (1977).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721-732 (1977).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- Tegtmeyer, P. *J. Virol.* **10**, 591-598 (1972).
- Redd, S. I., Stark, G. R. & Alwine, J. C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3083-3087 (1976).

- Khouri, G. & May, E. J. *Virology* **23**, 167-176 (1977).
- Leder, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 6187-6191 (1978).
- Khouri, G., Alwine, J., Goldman, N., Gruss, P. & Gilbert, W. *J. Virol.* (in the press).
- Konkel, D. A., Tilghmann, S. M. & Leder, P. *Cell* **15**, 1125-1132 (1978).
- Curtis, P. J. & Weissmann, C. *J. molec. Biol.* **106**, 1061-1075 (1976).
- Ross, J. *J. molec. Biol.* **106**, 403-420 (1976).
- Nordstrom, J. L., Roop, D. R., Tsai, M.-J. & O'Malley, B. W. *Nature* **278**, 328-331 (1979).
- Ryffel, G. V., Wyler, T., Muellener, D. B. & Weber, R. *Cell* **19**, 53-61 (1980).
- Lerner, M. R. & Steitz, J. A. *Proc. natn. Acad. Sci. U.S.A.* **11**, 5495-5499 (1979).
- Reddy, V. B. *et al. Science* **200**, 494-502 (1978).
- Fiers, W. *et al. Nature* **273**, 113-120 (1978).
- Sgarbella, V. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3389-3393 (1972).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Englund, P. T. *J. biol. Chem.* **246**, 3269-3276 (1971).
- Frost, E. & Williams, J. *Virology* **91**, 39-50 (1978).

LETTERS

Eddy diffusion coefficient for the atmosphere of Venus from radio scintillation measurements

Richard Woo

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91103

Akira Ishimaru

Department of Electrical Engineering, University of Washington, Seattle, Washington 98195

Estimates of the vertical mass eddy diffusion coefficient K are necessary to understand the dynamics and photochemistry of a planetary atmosphere. Most previous estimates for the atmosphere of Venus¹⁻¹⁰, pertain to the upper atmosphere. In the lower atmosphere (≤ 80 km), information on small-scale (≤ 1 km) temperature fluctuations has been provided by radio scintillations observed during radio occultation by the atmosphere¹¹⁻¹⁸. Here we relate the vertical mass eddy diffusion coefficient to these radio scintillation observations so that reliable estimates can be obtained near the tropopause level of 60 km for the first time. Using the recent Pioneer Venus radio occultation measurements¹⁶, and assuming that the turbulence is generated by shear instability, we find that $K = 4 \times 10^4 \text{ cm}^2 \text{ s}^{-1}$, the energy dissipation rate $\epsilon = 20 \text{ cm}^2 \text{ s}^{-1}$, and the dissipation rate of temperature fluctuations $N_\theta = 10^{-3} \text{ K}^2 \text{ s}^{-1}$. These results indicate that small-scale turbulence probably plays a major part in vertical transport near the tropopause of Venus.

Two regions of intense small-scale temperature fluctuations located near 45 and 60 km and coinciding with highly statically stable regions of the atmosphere of Venus have been identified by the radio scintillation measurements^{14,15}. We will study the upper region of turbulence. The turbulent eddies are anisotropic and elongated in the horizontal direction, and their wavenumber spectrum is approximately Kolmogorov. The Pioneer Venus orbit 18 entrance radio occultation measurements have been analysed in detail both for the mean temperature profile¹⁹ and for fluctuations in the upper region of turbulence¹⁶ situated above the tropopause. The estimate of structure constant c_T for the temperature fluctuations in the vertical direction is $0.15 \text{ K m}^{-1/3}$. This estimate is reliable because: first, the wave propagation theory used to interpret the radio scintillations is the most comprehensive to date²⁰; second, contrary to earlier results, the radio scintillations observed by entry probes and orbiting or fly-by spacecraft now show good agreement with each other²¹; and third, temperature fluctuations inferred from the radio measurements agree closely with the Pioneer Venus *in situ* results¹⁶. The latter were obtained by estimating the deviations from a linear fit to the temperature profile.

For the Kolmogorov spectrum, the structure constant c_T is related to the energy dissipation rate ϵ and the dissipation rate of temperature fluctuations N_θ by^{22,23}

$$c_T^2 = b N_\theta \epsilon^{-1/3} \quad (1)$$

where b is an empirical constant which is ~ 2.8 (ref. 24). The eddy diffusion coefficients (diffusivities) for momentum and heat, K_m and K_h respectively, are given by²⁵⁻²⁷

$$\epsilon = K_m \left(\frac{\partial u}{\partial z} \right)^2 (1 - R_i) \quad (2)$$

$$N_\theta = K_h \left(\frac{\partial \theta}{\partial z} \right)^2 \quad (3)$$

where R_i is the flux Richardson number, θ is the potential temperature, u is the horizontal wind speed and z is the vertical direction. K_m and K_h are also called eddy viscosity and eddy conductivity respectively. The flux Richardson number is related to the gradient Richardson number Ri by

$$R_i = \frac{K_h}{K_m} Ri \quad (4)$$

and

$$Ri = \frac{g \left[\frac{\partial \theta}{\partial z} / \left(\frac{\partial u}{\partial z} \right)^2 \right]}{\quad} \quad (5)$$

where the acceleration due to gravity g is 8.9 m s^{-2} for Venus. The eddy Prandtl number which is K_m/K_h has been found to be close to unity in the Earth's atmosphere^{23,28}. If we assume this to be the case of Venus as well and call $K = K_m = K_h$ the eddy diffusion coefficient, we obtain from equations (1)–(5)

$$c_T^2 = b \frac{\theta}{g} \frac{\partial \theta}{\partial z} \epsilon^{2/3} \frac{Ri}{1 - Ri} \quad (6)$$

a result which has been presented by Ottersten²⁹, and

$$K = \left[\frac{g(1 - Ri)}{Ri} \right]^{1/2} \frac{c_T^3}{b^{3/2} (\partial \theta / \partial z)^{5/2}} \quad (7)$$

Equation (6) indicates that the scintillations are caused by temperature fluctuations (in the case of Venus, the effect of water vapour fluctuations is negligible) which are directly related to both the mean vertical gradient of potential temperature and ϵ , a measure of the intensity of the velocity fluctuations. For a neutrally stable atmosphere, even if large velocity fluctuations are present, no scintillations would be detected as $\partial \theta / \partial z = 0$. Thus, although scintillations were not detected in the middle cloud region of the Venus atmosphere near 50 km (ref. 16), it cannot be concluded that mechanical turbulence is absent because the stability of the atmosphere is very close to neutral³⁰.

The fact that the scintillations are observed in regions of high static stability where high wind shears are also present suggests that the turbulence is produced by wind-shear instability¹⁶. At 60 km, the Pioneer Venus measurements^{30,31} of θ and u yield $Ri = 1.7$ ($\partial \theta / \partial z = 1.6 \times 10^{-2} \text{ K m}^{-1}$ and $\partial u / \partial z = 10^{-2} \text{ s}^{-1}$). Although Ri is small, it falls short of 0.25 which is the critical value for initiation of dynamic-shear instability indicated by both theoretical and experimental studies³². This is, however, not surprising as the Pioneer Venus measurements of temperature and wind velocity lack the spatial resolution to measure the local gradient Richardson number adequately. In the case of the terrestrial free atmosphere, estimates of Ri decrease substantially with decrease in scale over which they are measured³³. Typically, $Ri = 0.25$ only in thin layers³⁴ and generally escapes detection unless the spatial resolution is of the order of 100 m. In analysing the scintillation observations of Venus we will therefore assume that $Ri = 0.25$.

Note that when $Ri = 0.25$ and $K_h = K_m$, equations (2)–(5) yield

$$\epsilon = 3 N_\theta^2 K_h \quad (8)$$

where $N = (g/\theta) (\partial \theta / \partial z)^{1/2}$ is the Brunt-Väisälä frequency. Equation (8) is the relationship used by Lilly *et al.*³⁵ to deduce K_h in the Earth's stratosphere.

Let us now estimate K for Venus at 60 km using equation (7). For $c_T = 0.15 \text{ K m}^{-1/3}$, $b = 2.8$, $Ri = 0.25$, $\partial u / \partial z = 2.6 \times 10^{-2} \text{ s}^{-1}$ and $\partial \theta / \partial z = 1.6 \times 10^{-2} \text{ K m}^{-1}$, we obtain $K = 4 \times 10^4 \text{ cm}^2 \text{ s}^{-1}$. In comparing the tropopause regions of the Earth^{28,35} and Venus

we find that the eddy diffusion coefficients are similar. Note that the assumed wind shear corresponding to $Ri = 0.25$ is $2.6 \times 10^{-2} \text{ s}^{-1}$, which is a factor of 2.6 higher than the observed value. These strong shears are not unusual in the Earth's free atmosphere and have been observed over altitude intervals of several hundred metres³⁶.

For $K = 4 \times 10^4 \text{ cm}^2 \text{ s}^{-1}$, equations (2) (3) give $\varepsilon = 20 \text{ cm}^2 \text{ s}^{-3}$ and $N_g = 10^{-3} \text{ K}^2 \text{ s}^{-1}$. The energy dissipation rate ε deduced from the Pioneer Venus radio scintillations can be compared with estimates obtained from the wind measurements. The latter can be obtained using the relationship²²⁻²⁵

$$\varepsilon = \frac{(\sigma_v)^3}{L} \quad (9)$$

where σ_v is the r.m.s. velocity and L is the characteristic spatial scale. In the case of the Pioneer Venus measurements³¹, we estimate that at 60 km $L \sim 5 \text{ km}$ and $\sigma_v \sim 2 \text{ m s}^{-1}$ which is consistent with the Venera 9–12 results^{37,38}. This leads to $\varepsilon = 18 \text{ cm}^2 \text{ s}^{-3}$ which agrees well with the radio results. Kerzhanovich *et al.*³⁸ have analysed the Venera 11 and 12 wind measurements and obtained $50\text{--}180 \text{ cm}^2 \text{ s}^{-3}$ for ε above 40 km. All these estimates of ε for Venus fall within the range measured near the Earth's tropopause using *in situ*^{28,35} and radar^{39,40} observations.

Turbulence in the upper troposphere and lower stratosphere of Venus seems to be similar to that found near the terrestrial tropopause which is due primarily to the Kelvin–Helmholtz instability^{41,42}. It is produced in the region of high wind shear⁴³ which exists above 40 km (ref. 31). Temperature fluctuations are generated in the highly statically stable regions near 45 and 60 km (ref. 30) and these in turn give rise to the radio scintillations that are observed. In the vicinity of 50 km (middle cloud region), the stability of the atmosphere is nearly neutral so that temperature fluctuations are substantially reduced and no scintillations occur. Covey and Schubert⁴⁴ have recently discussed the eddy diffusion coefficient for the middle cloud region near the sub-solar point.

Turbulence near the terrestrial tropopause generally takes the form of thin horizontal layers³⁹⁻⁴¹. The radio occultation measurements of Venus do not have enough resolution to determine whether the patchiness of terrestrial turbulence also exists in the atmosphere of Venus. However, because the radio scintillation observation is a path-integrated measurement, the estimate of eddy diffusion coefficient obtained in this study is representative of the average atmosphere at 60 km. The equations which were used to estimate K are based on isotropic turbulence and strictly speaking do not apply to the anisotropic turbulence probed by Pioneer Venus. On the other hand, the structure constant c_T used to deduce K is that in the vertical direction and accounts for the fact that the turbulence is stronger in the vertical than in the horizontal direction.

We have found that K is $\sim 4 \times 10^4 \text{ cm}^2 \text{ s}^{-1}$ at 60 km, and believe that this is a good order-of-magnitude estimate. Note that this result is close to the value often assumed in dynamical studies of the atmosphere of Venus^{46,47}. Although the scales measured by scintillations are small¹⁶ ($\leq 1 \text{ km}$), they are substantially larger than those in the terrestrial troposphere that have been measured and studied extensively with radar^{39,40} ($\leq 10 \text{ m}$). The fact that the eddy diffusion coefficient estimated from the radio scintillations is consistent with that expected on the basis of vertical cloud-particle distribution⁴ and photochemical¹⁰ considerations, implies that small-scale turbulence is probably the dominant mechanism for vertical transport near the tropopause ($\sim 60 \text{ km}$) in the Venus atmosphere. This is not the case in the Earth's atmosphere^{39,48}.

Clearly, the radio scintillation technique provides a useful means for determining the eddy diffusion coefficient in regions of the atmosphere where measurements have not yet been made. These radio measurements are applicable to the outer planets, and would yield not only details of the small-scale turbulence in the lower atmosphere, but information on the static stability of the atmosphere as well.

We thank J. W. Armstrong, E. M. Dewan and R. E. Young for useful discussions. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract NAS7-100, sponsored by NASA.

Received 30 September; accepted 26 November 1980.

- Liu, S. C. & Donahue, T. M. *Icarus* **24**, 148–156 (1975).
- Sze, N. D. & McElroy, M. B. *Planet. Space Sci.* **23**, 763–786 (1975).
- Hunten, D. in *Atmospheres of Earth and the Planets* (ed. McCormac, B. M.) 59–72 (Reidel, Dordrecht, Holland, 1975).
- Prinn, R. *J. Atmos. Sci.* **31**, 1691–1697 (1974).
- Prinn, R. *J. Atmos. Sci.* **32**, 1237–1247 (1975).
- Belton, M. J. S., Smith, G. R., Schubert, G. & Del Genio, A. D. *J. Atmos. Sci.* **33**, 1394–1417 (1976).
- Kumar, S., Hunten, D. M. & Broadfoot, A. L. *Planet. Space Sci.* **26**, 1063–1075 (1978).
- von Zahn, U., Fricke, K. H., Hoffman, H. J. & Pelka, K. *Geophys. Res. Lett.* **6**, 337–340 (1979).
- von Zahn, U. *et al. J. Geophys. Res.* **85**, 7829–7840 (1980).
- Winick, J. R. & Stewart, A. I. *J. Geophys. Res.* **85**, 7849–7860 (1980).
- Gurvich, A. S. *Doklady. Akad. Nauk SSSR* **5**, 1172–1178 (1969).
- Golitsyn, G. S. & Gurvich, A. S. *J. Atmos. Sci.* **28**, 138–140 (1971).
- de Wolf, D. A. *J. Geophys. Res.* **76**, 3154–3158 (1971).
- Woo, R., Ishimaru, A. & Kendall, W. B. *J. Atmos. Sci.* **31**, 1698–1706 (1974).
- Woo, R. *J. Atmos. Sci.* **32**, 1084–1090 (1975).
- Woo, R., Armstrong, J. W. & Ishimaru, A. *J. Geophys. Res.* **85**, 8031–8038 (1980).
- Timofeeva, T. S., Yakovlev, O. I. & Efimov, A. I. *Cosmic Res.* **16**, 285–293 (1978).
- Kolosov, M. A. *et al. Acta astr.* **7**, 219–234 (1980).
- Kliore, A. J. & Patel, I. R. *J. Geophys. Res.* **85**, 7957–7962 (1980).
- Woo, R., Ishimaru, A. & Yang, F. C. *Radio Sci.* **15**, 695–703 (1980).
- Woo, R., Armstrong, J. W. & Kendall, W. B. *Science* **205**, 204–206 (1979).
- Tatarskii, V. I. *The Effects of the Turbulent Atmosphere on Wave Propagation* (National Technical Information Service, Springfield, 1971).
- Lumley, J. L. & Panofsky, H. A. *The Structure of Atmospheric Turbulence* (Wiley, New York, 1964).
- Monin, A. S. & Yaglom, A. J. *Statistical Fluid Mechanics: Mechanics of Turbulence* Vol. 2 (MIT, Cambridge, Massachusetts, 1975).
- Ishimaru, A. *Wave Propagation and Scattering in Random Media* Vol. 2 (Academic, New York, 1978).
- Vinnichenko, N. K., Pinus, N. Z., Shmeter, S. M. & Shur, G. N. *Turbulence in the Free Atmosphere* (Consultants Bureau, New York, 1973).
- Dutton, J. A. & Panofsky, H. A. *Science* **167**, 937–944 (1970).
- Kennedy, P. J. & Shapiro, M. A. *J. Atmos. Sci.* **37**, 986–993 (1980).
- Ottersten, H. *NOAA Tech. Rep. ERL 243-WPL* (1972).
- Seiff, A. *et al. J. Geophys. Res.* **85**, 7903–7933 (1980).
- Counselman, C. C. III, Gourevitch, S. A., King, R. W. & Loriot, G. B. *J. Geophys. Res.* **85**, 8026–8030 (1980).
- Woods, J. D. *Radio Sci.* **4**, 1289–1298 (1969).
- Reiter, E. R. & Lester, P. F. *Arch. Met. Geophys. Bioklim.* **A17**, 1–7 (1968).
- Buffon, J. W. *J. Atmos. Sci.* **30**, 83–87 (1973).
- Lilly, D. K., Waco, D. E. & Adelfang, S. I. *J. appl. Met.* **13**, 488–493 (1974).
- Browning, K. A., James, P. K., Parkes, D. M., Rowley, C. & Whyman, A. J. *Nature* **271**, 529–531 (1978).
- Marov, M. Ya. *Astr. Astrophys.* **16**, 141–169 (1978).
- Kerzhanovich, V. V., Makarov, Yu. F., Marov, M. Ya., Rozhdestvenkiy, M. K. & Sorokin, V. P. *Academy of Sciences USSR Rep.* (transl. NASA TM-75726, 1979).
- Crane, R. K. *Radio Sci.* **15**, 177–193 (1980).
- Gage, K. S., Green, J. L. & van Zandt, T. E. *Radio Sci.* **15**, 407–416 (1980).
- Dutton, J. A. *Rev. Geophys. Space Phys.* **9**, 613–657 (1971).
- Thorpe, S. A. *Boundary-Layer Met.* **5**, 95–119 (1973).
- Crisp, D. & Young, A. T. *Icarus* **35**, 182–188 (1978).
- Covey, C. & Schubert, G. *Nature* (in the press).
- Dewan, E. M. *Science* **204**, 832–835 (1979).
- Goody, R. M. & Robinson, A. R. *Astrophys. J.* **146**, 339–355 (1966).
- Young, R. E. & Schubert, G. *Planet Space Sci.* **21**, 1563–1580 (1973).
- Reiter, E. R. *Rev. Geophys. Space Phys.* **13**, 459–474 (1975).

Optical observations of OV–236

D. Wills & Beverley J. Wills

McDonald Observatory, Department of Astronomy,
University of Texas, Austin, Texas 78712

Dent and Balonek¹ recently reported a large outburst in the radio source OV–236 (=1921–293). We have made optical observations of the associated optical object, and report here that it has a high degree of linear polarization, optically variability of $\sim 1 \text{ mag}$ in 25 days, and an emission-line spectrum at a redshift $z = 0.3525$. It thus belongs to the small class of extragalactic objects whose properties are intermediate between those of 'classical' QSOs and BL Lac objects.

There has been considerable confusion about the correct optical identification of OV–236. Dent and Balonek¹ refer to the red stellar object suggested in ref. 2, and give an optical

position, apparently from ref. 3, but neither is correct. Accurate radio positions that agree within 0.2 arcs have been measured^{4,5}, but the red object noted in ref. 2 is 37 arcs away, and is therefore definitely excluded as the optical identification. The optical object is correctly identified in ref. 3 and in ref. 5, where its radio and optical position agree within ~ 1 arcs. We have also measured the 1950.0 optical position from glass plate copies of the Palomar Sky Survey, and find RA = 19 h 21 min 42.26 s, dec = $-29^{\circ}20'27.0''$, with r.m.s. uncertainties of ~ 0.4 arcs; this position is within 0.3 arcs of that given in ref. 5. The optical position given in ref. 3, and repeated in ref. 1, is probably a misprint—there is no object visible there on the Sky Survey plates. The correct identification appears stellar, and about 18 mag, on two pairs of Sky Survey plates (epochs 1954.72 and 1958.53). It may be slightly brighter, and again appears stellar, on the SRC IIIa-J Schmidt plate J2597 (epoch 1976.71), which reaches fainter magnitudes and has better resolution. The object appears to be red on the Sky Survey plates, at least in part because of galactic extinction—its galactic coordinates are $l = 9.3^{\circ}$, $b = -19.6^{\circ}$. It has a history of optical variability⁵. Unfortunately, the radio outburst described in ref. 1 occurred at a time of year when the object was becoming inaccessible to ground-based optical telescopes.

We obtained polarimetry of the object on 12.4 June 1980 UT, using the McDonald Observatory 2.1-m Struve reflector, and the polarimeter described in ref. 6. Visually, the object was estimated to be ~ 1 mag brighter than on the Sky Survey plates, and equal in brightness to the star 19 arcs following; that is, ~ 17 mag. The degree of linear polarization was found to be $8.0 \pm 0.5\%$ r.m.s., in position angle $149 \pm 2^{\circ}$. The galactic reddening that we adopt below implies a contribution of $< 1\%$ polarization due to the interstellar medium⁷.

Spectrophotometry was obtained between 3,900 and 6,900 Å on 14 July 1980 UT, and between 3,400 and 5,700 Å on 8 August 1980, with ~ 7 Å FWHM resolution, using the McDonald 2.7-m NASA reflector and the instrument described in ref. 8. Figure 1 shows the spectra, calibrated by observations of standard stars to an accuracy of $\sim 10\%$ r.m.s. The difference in continuum levels indicates a fading of nearly 1 mag during the 25 days separating the two sets of observations (J. T. Pollock found a decrease of ~ 0.7 mag in m_{pg} between 16 July and 7 August).

The first spectrum shows two narrow (~ 15 Å FWHM) emission lines at 5,042.5 and 6,770.5 Å, which can be unambiguously identified as [O II] and [O III] $\lambda\lambda 3,727, 5,007$ at an observed redshift $z = 0.3525 \pm 0.0005$ r.m.s. This spectrum also shows an emission line at the expected wavelength of [O III] $\lambda 4,959$, but no clear sign of H β . The second spectrum again shows [O II], at 5,042.9 Å, and a narrow feature at the position of Mg II $\lambda 2,798$. Although the spectrum is noisy near the position of Mg II, the width of the feature is consistent with those of the forbidden lines. Neither spectrum shows evidence of absorption lines that might indicate the presence of a galaxy.

In the first spectrum the rest-frame equivalent widths of the [O II] and [O III] lines are 3.7 and 9.2 Å, respectively, while in the second one the values for [O II] and Mg II are 9.0 and 4.1 Å, respectively (with r.m.s. uncertainties of typically 10%). The large and significant change in the equivalent width of [O II] is independent evidence for a continuum level change, because the intensity of this line appeared to change by only $\sim 25\%$ between the two observations, a difference that is within the uncertainties in the absolute levels of the two spectra ($\sim 10\%$ r.m.s. per observation). The [O II] emission arises in a large low-density region and is not expected to change significantly on a time scale of 25 days (18.5 days at the object). Based on the latter assumption, the fainter spectrum in Fig. 1 has been scaled by a factor of 0.734 so as to equalize the intensities of [O II].

The two spectra in Fig. 1 have the same shapes, with spectral indices $\alpha = 1.3$ and $\alpha = 2.4$ over 5,300–6,750 Å and 3,800–5,300 Å, respectively, after applying a correction for galactic reddening, using $A_V = 0.5$ mag (see for example, ref. 10), and defining the spectral index α by $f_{\nu} \propto \nu^{-\alpha}$. The steepening

below 3,600 Å may have been exaggerated by the use of a 2.8 (NS) by 4.5 (EW) arcs aperture and the large zenith distance ($\sim 60^{\circ}$) of the object, but we do not believe that these factors contribute to the spectrum curvature between 3,800 and 6,750 Å (this is supported by the good agreement between the spectral shapes in the wavelength region that is common to both spectra). The difference between the two spectra obtained in July and August in the region of overlap is well-represented by a power-law spectrum having $\alpha = 2.1 \pm 0.1$ r.m.s. ($\alpha = 2.9$ before de-reddening); power-law spectra were also derived for the variable components in the variable QSOs 0736 + 01 and 3C345 (ref. 9).

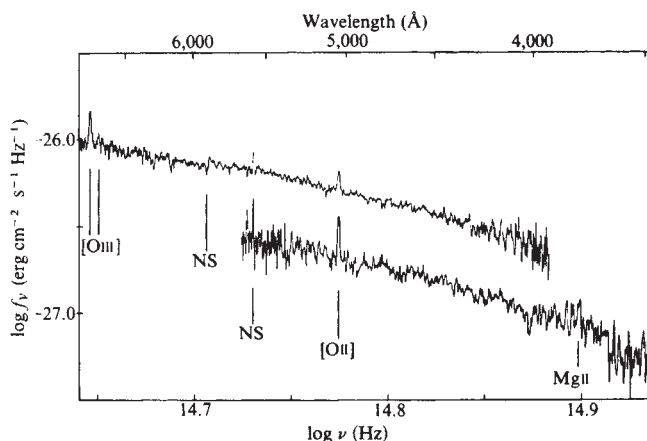


Fig. 1 Spectra of the optical object associated with OV-236, obtained with the McDonald Observatory IDS. The difference in continuum level represents real variability of the object during the 25 days that separate the two observations. Emission lines of Mg II $\lambda 2,798$, [O II] $\lambda 3,727$ and [O III] $\lambda\lambda 4,959, 5,007$ are marked. NS indicates the positions of incompletely cancelled night sky lines ($\lambda\lambda 5,577, 5,892$).

The high degree of linear polarization, the neutral colour and the emission-line spectrum form an unusual, but not unique, combination, placing OV-236 intermediate in properties between the BL Lac objects and QSOs. Most of the objects that are strongly variable at radio and optical frequencies, and that have high ($> 5\%$) optical polarization, have optical spectra that show no emission lines when observed with the sensitivity and wavelength resolution appropriate to the present observations. Had the object been 1.5 mag brighter than in July (as it was in mid-1979, ref. 5), with the emission lines retaining their intrinsic strength⁹, the lines would have been barely detectable (1–3 Å equivalent width), and we would probably have classified OV-236 as a classical BL Lac object.

Perhaps the object that is most similar to OV-236 is 1400 + 162 (refs 11, 12), which is neutral in colour, has high optical polarization, and even weaker emission lines than those in OV-236 (from Fig. 2 of ref. 11 we estimate the rest-frame equivalent width of [O II] $\lambda 3,727$ in 1400 + 162 to be 1.0 Å, and the intrinsic luminosity of the [O III] $\lambda 5,007$ line in OV-236 is about 45 times that in 1400 + 162 (ref. 12)). No radio maps of OV-236 are yet available, but the fact that its 408 MHz flux density¹³ is close to the pre-eruption value at high frequencies¹ suggests that the source does not possess extended components with steep radio spectra like those of 1400 + 162 (ref. 11).

Other possibly similar objects are 0752 + 258, which has a high degree of linear polarization^{14,15}, although much stronger emission lines^{16,17}, and 3C446, with its history of radio and optical variability, its high polarization and its rather weak emission lines¹⁸. Further observations of this small group of objects may eventually help to elucidate the connection between BL Lac objects and QSOs.

We thank J. T. Pollock for information about the recent optical variability of OV-236, and the NSF for support.

Received 6 August; accepted 15 November 1980.

1. Dent, W. A. & Balonek, T. J. *Nature* **283**, 747 (1980).
2. Radovich, M. M. & Kraus, J. D. *Astr. J.* **76**, 683 (1971).
3. Peterson, B. A., Bolton, J. G. & Shimmins, A. J. *Astrophys. Lett.* **15**, 109 (1973).
4. Gubbay, J. S. *Astr. J.* **83**, 697 (1978).
5. Donovan, F. F. *et al. Publ. Astrophys. Space Phys.* (in the press).
6. Breger, M. *Astrophys. J.* **233**, 97 (1979).
7. Serkowski, K., Mathewson, D. S. & Ford, V. L. *Astrophys. J.* **196**, 261 (1975).
8. Wills, B. J. & Wills, D. *Astrophys. J.* **238**, 1 (1980).
9. Netzer, H., Wills, B. J., Uomoto, A. K., Rybski, P. M. & Tull, R. G. *Astrophys. J. Lett.* **232**, L155 (1979).
10. Schmidt, M. *Astrophys. J.* **151**, 393 (1968).
11. Baldwin, J. A. *et al. Astrophys. J.* **215**, 408 (1977).
12. Miller, J. S., French, H. B. & Hawley, S. A. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 176 (University of Pittsburgh Press, 1978).
13. Large, M. L., Mills, B. Y., Little, A. G., Crawford, D. F. & Sutton, J. M. *Mon. Not. R. astr. Soc.* (in the press).
14. Stockman, H. S. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 149 (University of Pittsburgh Press, 1978).
15. Wills, D., Wills, B. J., Breger, M. & Hsu, J.-C. *Astr. J.* (in the press).
16. Wills, D. & Wills, B. J. *Astrophys. J. Suppl.* **31**, 143 (1976).
17. de Vaucouleurs, G., de Vaucouleurs, A. & Nieto, J.-L. *Astr. J.* **84**, 1811 (1979).
18. Miller, J. S., & French, H. B. *Pittsburgh Conf. on BL Lac Objects* (ed. Wolfe, A. M.) 228 (University of Pittsburgh Press, 1978).

Excited ozone is a possible source of atmospheric N₂O

Sheo S. Prasad

Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Drive, Pasadena, California 91109

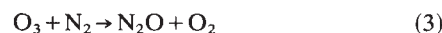
In the formation of ozone from O and O₂ in the ground state, the products of the reaction may include electronically excited forms of ozone (O₃^{*}) which are spectroscopically recognizable in pulsed radiolysis and flash photolysis experiments¹⁻⁴. Indeed, it has been suggested that in the three-body recombination reaction



there is a 62% chance that the ozone product will be in an excited form. Although the identity of these excited precursors of ozone is not known, excited ³B₂, ³A₂ and ¹A₂ states of bent ozone and the cyclic ozone have been suggested⁵⁻⁷, and the existence of electronically excited states of bent ozone below its dissociation limit has been corroborated by large-angle electron scattering experiments⁸. Thus, more than one form of internally excited ozone precursor may be involved. It has been estimated⁹ that $\Sigma n(\text{O}_3^*) \approx (100-1,000) \cdot n(\text{O}^1\text{D})$ in the stratosphere. Consequently, even with moderate reactivity, O₃^{*} may be chemically significant in the stratosphere. We now hypothesize that some of the internally excited ozone, probably that which is electronically excited, may be a potential source of N₂O through the gas-phase reaction (2a),



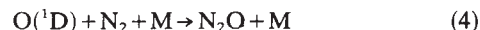
Previously, N₂O was considered¹⁰⁻¹² to form by the reaction



involving unexcited ozone, for which a rate coefficient $k_3 = 5 \times 10^{-28} \text{ cm}^3 \text{ s}^{-1}$ at 292 K has been reported¹³. This rate constant was estimated by mixing nitrogen and ozone in an absorption tube, measuring ozone disappearance from the absorption in Chappuis band and nitrous oxide formation from the intensity in the 7.8 μm band. However, such a slow reaction between two quite stable molecules is unlikely to proceed

without substantial activation energy. In the colder upper atmosphere, therefore, k_3 will probably be much smaller than $5 \times 10^{-28} \text{ cm}^3 \text{ s}^{-1}$. Hence reaction (3) is not expected to be a significant atmospheric source of N₂O. On the other hand, reaction (2a) seems more likely in view of Goody and Walshaw's results¹³, because there are many known instances where internal excitation (either vibrational or electronic) in one reactant significantly enhances the reaction.

The wavelength dependence of the quantum yield for N₂O formation, $\phi(\text{N}_2\text{O})$, from photolysis of dilute solutions of O₃ in liquid N₂ at 77 K (ref. 14), and the wavelength dependence of O(¹D) quantum yield in gas-phase photolysis of O₃ (refs 15-19) may be interpreted to suggest the possibility of reaction (2a). Relevant experimental data are presented in Fig. 1, where $\phi(\text{N}_2\text{O})$ obtained by DeMore and Raper¹⁴ is shown with $\phi(\text{O}^1\text{D})$ obtained in gas-phase experiments. Over most of the wavelengths used in DeMore and Raper's experiment, $\phi(\text{N}_2\text{O})$ is synonymous with $\phi(\text{O}^1\text{D})$ in ozone photolysis. Thus, $\phi(\text{N}_2\text{O})$ in this experiment begins to decrease rapidly at about 300 nm in agreement with the behaviour of $\phi(\text{O}^1\text{D})$ in gas-phase experiments. However, $\phi(\text{N}_2\text{O})$ in DeMore and Raper's experiment is perceptible even at 334 nm where $\phi(\text{O}^1\text{D})$ is most probably zero according to other experiments. This implies that an additional, low activation energy mechanism of N₂O formation exists, which comes into prominence in DeMore and Raper's experiments at longer wavelengths where $\phi(\text{O}^1\text{D})$ vanishes and the reaction



becomes inoperative. We conjecture that reaction (2a) may be this new mechanism. We therefore augmented DeMore and Raper's reaction scheme by including our reactions (1) and (2), and a new expression for N₂O quantum yield is derived:

$$\phi(\text{N}_2\text{O}/\lambda) \approx \frac{\phi_1(\lambda)}{1 + k_6/k_4} + 0.62\phi_2(\lambda)f_1f_2 \quad (5)$$

where 0.62 is the probability of internal excitation in nascent ozone as explained before; f_1 is that fraction of internally excited nascent ozone which could produce N₂O; $f_2 = k_{2a}/k_2$ applicable in their case; ϕ_1 is the quantum yield for O(¹D); ϕ_2 is the same for O(³P) and k_6 is the rate coefficient for the reaction



At $\lambda < 300 \text{ nm}$, $\phi_1 = 1$ and $\phi_2 = 0$, and at 334 nm, $\phi_1 = 0$ and $\phi_2 = 1$, so that $f_1f_2 \approx 4.2 \times 10^{-3}$ on the basis of the observed N₂O quantum yields in DeMore and Raper's experiments¹⁴. On the other hand, failure to detect N₂O, clearly above the background

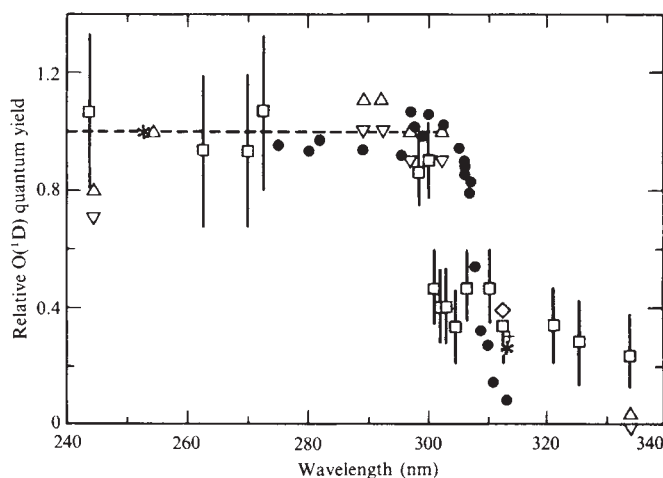


Fig. 1 Relative O(¹D) quantum yields at various wavelengths in the photolysis of ozone. □, Ref. 14; ●, ref. 15; ▽, ref. 16; +, ref. 17; ◇, ref. 18; *, ref. 19. DeMore and Raper¹⁴ report $\phi(\text{N}_2\text{O})$ which is synonymous with $\phi(\text{O}^1\text{D})$ over most of the wavelengths used by these authors.

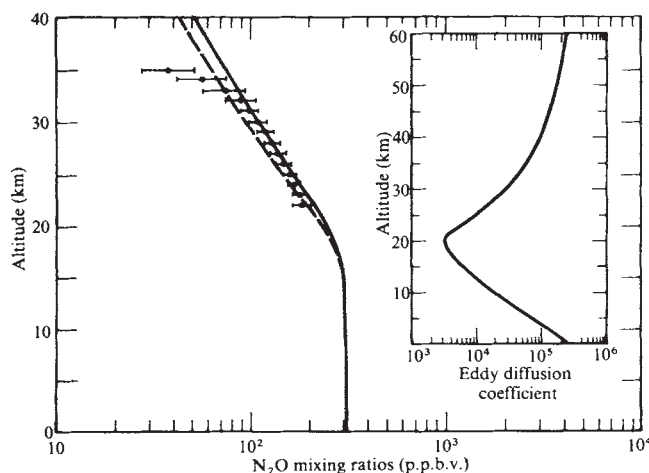


Fig. 2 Theoretical N_2O mixing ratios. Solid line, with the new N_2O source involving excited ozone; dashed line, without this source; $\bigcirc$ represent atmospheric measurements of Farmer *et al.*²³. The inset shows the profile of eddy-diffusion coefficients used in this study.

level, in the experiments of Simonaitis *et al.*²⁰ places an approximate upper limit of $f_1 f_2 \approx 10^{-4} - 10^{-5}$ in their experiments conducted at 292 K. Considering widely different experimental conditions and the uncertainties, the difference between the two estimates of $f_1 f_2$ is quite understandable, for example, in terms of possible inverse temperature dependence of $f_1 f_2$. However, because the present study is exploratory we have ignored details, such as the temperature dependence, and have arbitrarily used $f_1 f_2 = 10^{-5}$. This choice is consistent with Simonaitis *et al.*'s experiment which was in the gas phase and at a temperature close to the stratospheric temperatures.

The production rate of N_2O through reaction (2a) may now be written as:

$$q(\text{N}_2\text{O}) = \frac{6.2 \times 10^{-6} k_1 n(\text{O}) n(\text{O}_2) n(\text{M})}{k_2 n(\text{M}) + 1/\tau} k_2 n(\text{N}_2) \quad (7)$$

$$\approx \frac{6.2 \times 10^{-6} (J_{\text{O}_3} + 2J_{\text{O}_2})}{k_2 n(\text{M}) + 1/\tau} k_2 n(\text{N}_2). \quad (8)$$

where $J_{\text{O}_3, \text{O}_2}$ is the number of O_3 (or O_2) photodissociated per second. Approximate equality between reactions (7) and (8) becomes increasingly invalid above about 60 km. But this is immaterial because O_3^* -related source of N_2O is very small at these higher altitudes. As a first approximation, $k_2 (= k_{2a} + k_{2b})$ may be assumed to be the same for all internally excited ozone, so that $k_2 \sim (7-0.6) \times 10^{-14} \text{ cm}^3 \text{ s}^{-1}$ on the basis of several experimental data¹⁻³. There are no direct experimental data on the lifetime, τ , of internally excited ozone against non-collisional destruction. We have therefore tried several estimates ranging from a few microseconds to a millisecond. Atmospheric concentrations and fluxes of N_2O corresponding to its production through reaction (2a) were then calculated. These calculations were made for the US Standard Atmosphere, 1976, using one-dimensional photochemical-dynamical modelling. We assumed that N_2O molecules were lost by photodissociation and by reaction with $\text{O}(\text{D})$. $\text{O}(\text{D})$ concentrations were obtained on the basis of local chemical equilibrium between its production from O_3 photodissociation and loss by quenching collisions with N_2 and O_2 . Ozone distribution needed for this and for estimating the solar UV attenuation was taken from US Standard Atmosphere. The empirical method of Shimazaki *et al.*²¹ was used to calculate the absorption of solar UV in the O_2 Schuman-Runge bands. We assumed a N_2O mixing ratio of 3.1×10^{-7} at the lower boundary²² and a zero flux at the upper boundary at 90 km. Calculations were done for a latitude of 30° .

The results in Fig. 2 were obtained by using $\tau = 9 \times 10^{-6} \text{ s}$, $k_2 = 2.0 \times 10^{-14} \text{ cm}^3 \text{ s}^{-1}$ for the excited ozone able to react with N_2 (the eddy-diffusion coefficient, k_{eddy} , profile used is shown as

an inset). Our small value of τ is not inconsistent with the experimental data of refs 1-4, because it applies to only a small fraction f_1 of excited ozone which may approach 10^{-5} if f_2 approaches unity. Measurements of the N_2O mixing ratio at 32°N are also shown in Fig. 2 for comparison with our theoretical results. Our theoretical profile using the assumed excited ozone source of N_2O clearly agrees with Farmer *et al.*'s²³ data better than the profile without this source, if we ignore the data points at 34 and 35 km which have significantly larger experimental error bars. This performance of our theory is, of course, dependent on various factors in the calculations, such as the profile of k_{eddy} , and the properties of excited ozone precursors. Nevertheless, with a reasonable choice of k_{eddy} , the assumed N_2O production from excited ozone has reproduced certain observed nitrous oxide mixing ratios as satisfactorily as the theory without this assumed new source. Nitrous oxide mixing ratios exhibit great variability. Naturally, other observed profiles of nitrous oxide mixing ratios may not agree as well with our new theoretical profile. Use of other k_{eddy} profiles, particularly those determined from observed N_2O mixing ratios and the assumption that there is no atmospheric source for this species, will also spoil the good performance of the new theory evident in Fig. 2. These, however, do not contradict the new theory. The present spread in measured N_2O mixing ratios and estimates of k_{eddy} profiles do not allow a conclusive choice between the new theory and the current theory which links atmospheric N_2O solely to microbiological activity on the Earth's surface. Experimental investigations of reaction (2a) are needed.

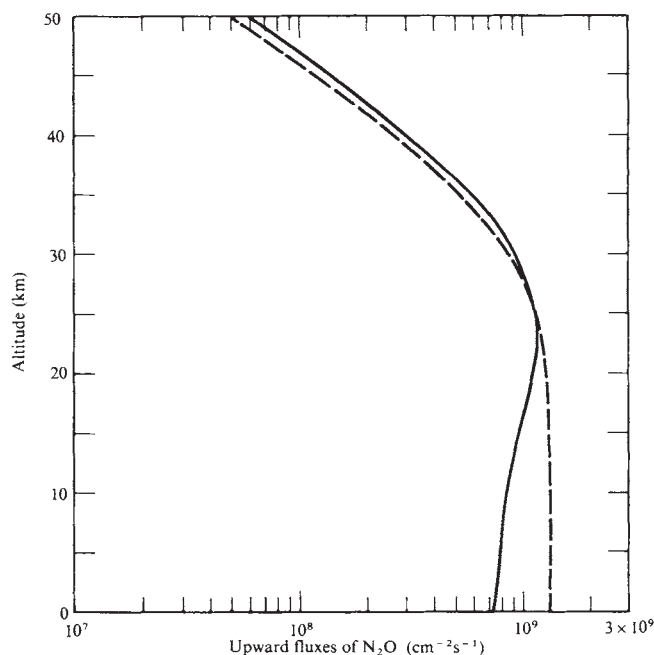


Fig. 3 Upward flux of N_2O corresponding to the profiles of nitrous oxide mixing ratio and eddy-diffusion coefficients shown in Fig. 2.

Fluxes of nitrous oxide in the two limits with and without the new source are presented in Fig. 3. The new source of N_2O implies that the net upward flux of N_2O due to microbiological activities at the Earth's surface may be significantly smaller than the currently accepted value of $(1.3-1.4) \times 10^9 \text{ N}_2\text{O}$ molecules $\text{cm}^{-2} \text{ s}^{-1}$. Thus, this new source tends to strengthen a recent conjecture²⁴ that the main impact of human beings on the stratospheric N_2O level may not be through the indirect mechanism of increased nitrogen fixation, but through the direct anthropogenic production²⁵ of N_2O . Besides, the assumed new source of N_2O has two more attributes: (1) it provides a possible new physical basis for latitudinal and temporal variations in N_2O , because upper atmospheric ozone itself has these

variations; and (2) this source could impart an additional stability to the ozone layer against external perturbations. Any perturbation which produces an increase in ozone will result in increased N_2O from the new source and thereby in increased NO_x . The increased NO_x will exert a stabilizing effect by arresting further increase in ozone. Similar arguments, but in the opposite sense, would apply to perturbations which produce initial decrease in the ozone layer.

This study represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract no. NAS7-100, sponsored by NASA.

Received 11 August; accepted 28 November 1980.

- Riley, J. F. & Cahill, R. W. *J. chem. Phys.* **52**, 3297 (1970).
- Bevan, P. L. T. & Johnson, G. R. A. *JCS Faraday Trans. I* **69**, 216 (1973).
- von Rosenberg, C. W. & Trainer, D. W. *J. chem. Phys.* **63**, 5348 (1980).
- Kleindienst, T., Locker, J. R. & Bair, E. J. *J. Photochem.* **12**, 67 (1980).
- Wright, P. C. *Planet. Space Sci.* **25**, 1177 (1977).
- Hay, P. J., Dunning, T. H. & Goddard, W. A. III *Chem. phys. Lett.* **23**, 457 (1973).
- Wright, J. S. *Can. J. Chem.* **51**, 139 (1973).
- Swanson, N. & Cellota, R. J. *Phys. Rev. Lett.* **35**, 783 (1975).
- Prasad, S. S. & Burton, P. G. *Planet. Space Sci.* **27**, 411 (1979).
- Bates, D. R. & Witherspoon, A. E. *Mon. Not. R. astr. Soc.* **112**, 101 (1952).
- Bates, D. R. & Hays, P. B. *Planet. Space Sci.* **15**, 189 (1967).
- Sinel'nikova, G. Ye. *Atmos. Ocean Phys.* **12**, 106 (1976).
- Goody, R. M. & Walshaw, C. D. Q. *Jl R. met. Soc.* **79**, 496 (1953).
- DeMore, W. B. & Raper, O. F. *J. chem. Phys.* **37**, 2048 (1962).
- Lin, C. L. & DeMore, W. B. *J. Photochem.* **2**, 161 (1973/1974).
- Jones, I. T. N. & Wayne, R. P. *Proc. R. Soc. A* **319**, 273 (1970).
- Kuis, S., Simonaitis, R. & Heicklen, J. J. *geophys. Res.* **80**, 1328 (1975).
- Simonaitis, R., Braslavsky, S., Heicklen, J. & Nicolet, M. *Chem. phys. Lett.* **19**, 601 (1973).
- DeMore, W. B. & Raper, O. F. *J. chem. Phys.* **44**, 1780 (1966).
- Simonaitis, R., Lissi, E. & Heicklen, J. J. *geophys. Res.* **77**, 4248 (1972).
- Shimazaki, T., Ogawa, T. & Farrell, B. C. *NASA Tech. Note TN D-8399* (NASA, Washington DC 1977).
- Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport* (US National Academy of Sciences, Washington DC 1979).
- Farmer, C. B., Raper, O. F., Robbins, B. D., Toth, R. A. & Muller, C. J. *geophys. Res.* **85**, 1621 (1980).
- Levy, H. II, Mahlman, J. D. & Moxim, W. J. *Geophys. Res. Lett.* **6**, 155 (1979).
- Weiss, R. F. & Craig, H. *Geophys. Res. Lett.* **3**, 751 (1976).

Flexural strength and porosity of cements

J. D. Birchall, A. J. Howard & K. Kendall

Technical Department, ICI Mond Division, PO Box 8, The Heath, Runcorn, Cheshire WA7 4QD, UK

A curious feature of hydraulic cements, such as those based on calcium silicate, calcium aluminate and calcium sulphate, is that they exhibit similarly low flexural strengths, typically between 3 and 10 MPa, despite their differing chemical composition, varying degrees of hydration and contrasting setting mechanisms¹⁻³. Because of these low strength values, unreinforced cements are never used in flexure or tension, and studies of cement strength are usually confined to compression. Those few studies which have considered flexural or tensile failure have concluded that hydraulic cements have an intrinsic maximum tensile strength of about 20 MPa^{4,5}. Here we demonstrate that the commonly observed flexural weakness of cement is due to the presence of large voids which are largely undetected by conventional methods of pore analysis such as gas adsorption and mercury porosimetry. The removal of such macro-defects results in flex strengths up to 70 MPa, despite the large volume of gel pores remaining in the material. These strength figures, comparable with those of sintered ceramics, have been achieved without the use of elevated pressures or temperatures, and without fibrous reinforcement.

Much previous work on the fracture of cement has been concerned with compressive failure and with attempts to relate total volume porosity and compressive strength. From such studies, empirical laws such as those of Feret, Abrams and Powers have been derived⁶⁻⁹. However, the mode of failure in compression is complex and the relationship of compressive strength to porosity is unlikely to be adequately described by a single equation¹⁰. On the other hand the tensile and flexural

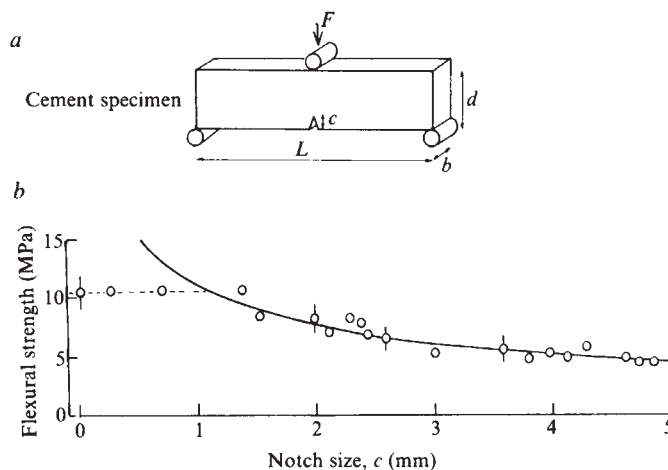


Fig. 1 a, The flexural testing arrangement. b, Results of flexural testing of notched, Ordinary Portland Cement paste.

strengths of brittle materials are more amenable to analysis¹¹⁻¹³ and in the Griffith equation,

$$\sigma = (ER/\pi c)^{1/2} \quad (1)$$

tensile cracking stress (σ) is related to an effective critical crack length (c) (which may be a pore); Young's modulus (E), and fracture surface energy (R) being taken as material constants.

The porosity of Portland cement paste is complex and involves a very wide range of pore sizes. Most of the porosity volume is associated with the hydrate gel structure (<100 nm) but larger porosity defects are present. There is a continuing debate as to which microstructural features limit the strength of cement pastes⁹. We have measured the flexural strengths of such pastes and related these through the Griffith equation to the sizes of the largest natural, or artificially-introduced 'pores' within the material. Finally, we have prepared cement paste specimens from which these macroscopic voids have been largely eliminated with a consequent, and dramatic, increase in strength.

The flexural strength testing arrangement is shown in Fig. 1a. Bars of cement paste were cast from a slurry containing 100 parts of Ordinary Portland Cement and 22 parts water, cured for 7 days and dried for a further 7 days. In some of the specimens, transverse notches of various depths were cut in the centre of one face using a diamond wheel. Notched and un-notched samples were then bent until cracking occurred. The data are presented in Fig. 1b as a Griffith plot of failure stress against

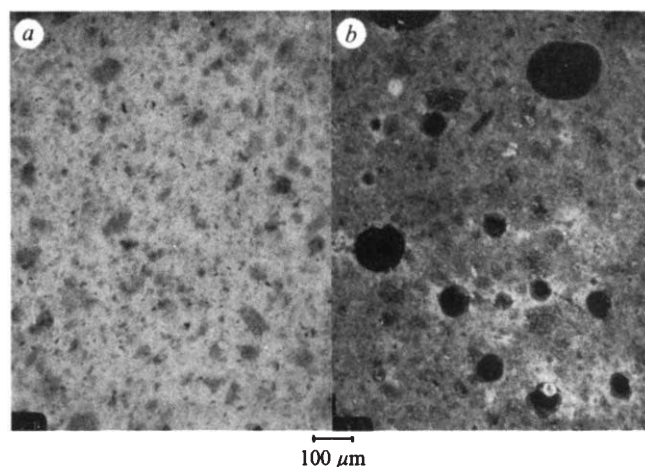


Fig. 2 Optical photomicrographs of polished cement paste sections. a, MDF cement. b, Ordinary Portland Cement paste. Voids are filled with Indian Ink to improve contrast, and appear black. The grey features are the residues of unhydrated cement grains. $\times 50$.

notch depth, the stress σ being calculated from the fracture force F using the relation $\sigma = 1.5 FL/bd^2$, as the flaw lengths were sufficiently small to avoid significant error¹⁴. The results gave a reasonable fit to the Griffith curve for notches exceeding 1.1 mm in length. Having calculated E from the flex measurements (20 GPa), the fracture energy $R = 19 \text{ J m}^{-2}$ was determined from equation (1). This value, typical of strong, brittle ceramics, indicates that the weakness of cement does not stem from low toughness. The results point to flaw size c as the cause of the weakness. For small notches, the strength was found to be essentially constant at $10.5 \pm 2 \text{ MPa}$, the same as un-notched samples. It was inferred that this deviation from the Griffith curve was caused by large natural flaws in the cement.

To investigate this, the porosity of the samples was measured both by mercury porosimetry¹⁵ and by quantitative microscopy¹⁶. Mercury porosimetry showed the pattern obtained by other workers^{17,18}, with a median pore diameter of 40–60 nm, a large proportion of gel pores (<7.5 nm diameter) and only 0.1% volume porosity above 15 μm diameter. In contrast, quantitative microscopy of polished sections of the cement demonstrated the presence of significant numbers of large holes, the

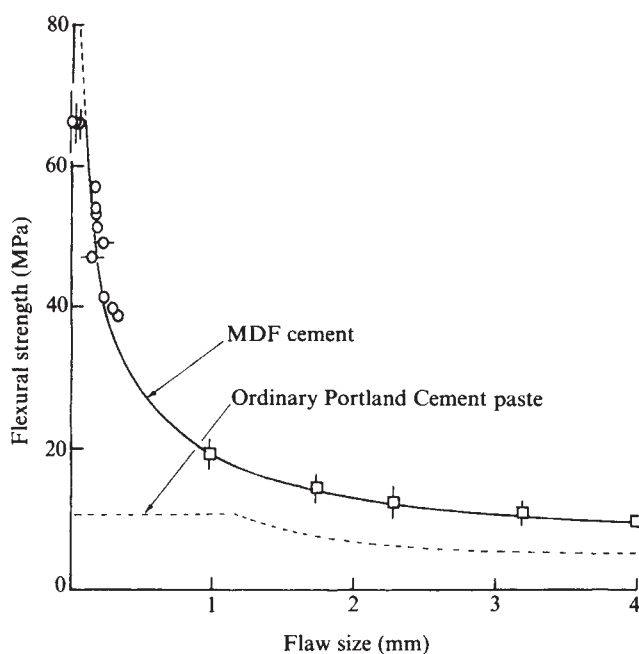


Fig. 3 Results of flexural testing of notched MDF cement (solid line) compared with Ordinary Portland Cement paste (dashed line). $\square$, Flaws introduced by diamond saw; $\circ$, flaws introduced by air bubbles or glass micro-balloons.

maximum hole size being measured at $0.83 \pm 0.17 \text{ mm}$ which agreed satisfactorily with the natural flaw size of $1.1 \pm 0.2 \text{ mm}$ obtained from the fracture measurements shown in Fig. 1. Furthermore, the volume associated with pores greater than 15 μm diameter was found to be about 5%, far greater than that given by mercury porosimetry, presumably because the mercury method cannot distinguish the pore diameter from the pore entrance diameter, or because the pores are effectively closed¹⁹.

Samples of cement paste were then prepared²⁰ in which the size of large included voids was reduced by a combination of rheological control and efficient mixing. Rheology was controlled, for example, by attention to particle size distribution and the use of rheological aids at low water/cement ratios (<0.2). Such methods reduced the size of the large voids whilst not necessarily reducing the total porosity volume as calculated from true and apparent density. A polished cross-section of a typical product, shown in Fig. 2a, illustrates the freedom from holes exceeding 15 μm in length, contrasting with the ordinary cement paste shown in Fig. 2b. Such preparations, (macro-defect-free (MDF) cement) were subjected to the bend test as

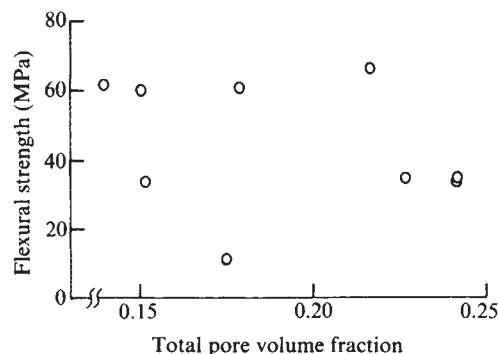


Fig. 4 The lack of correlation of flexural strength of cement paste with total pore volume. (Measured by mercury displacement and helium pycnometry.)

previously described. Macrodefects were artificially re-introduced as indicated in Fig. 3 which presents the results together with the data for the ordinary cement samples.

The results fitted the Griffith theory reasonably well but were different in two respects from the original cement. First, the strength at any given flaw size was higher. This was explained by the higher elastic modulus of MDF cement ($\sim 40 \text{ GPa}$) and its higher fracture surface energy of 30 J m^{-2} . The second difference was that MDF cement followed the Griffith curve to much higher stresses and lower flaw sizes than ordinary cement. When the largest visible holes were $\leq 90 \mu\text{m}$ in length (as measured by quantitative microscopy) the samples ceased to fit the Griffith curve but gave the same strength ($66 \pm 5 \text{ MPa}$) as material containing no holes visible in the optical microscope. Presumably, some flaws other than visible holes, perhaps grain interfaces, dominated at this juncture.

The experiments demonstrate that Griffith theory applies adequately to ordinary cement pastes and that the low flexural strengths normally found are associated with macroscopic flaws rather than microscopic or mineralogical features of the gel. It follows that total pore volume measurements will not correlate with strength except by accident, when pore size distribution remains constant over a series of samples. Figure 4 shows the lack of correlation of total porosity and strength in our experiments. Any attempts to relate flexural strength to gel microstructure, or to seek improved strength by altering microstructure will need to recognize the dominant role of large defects arising during preparation and not necessarily detected by conventional porosimetry.

We have shown that, by removing macroscopic flaws during preparation, cement pastes can be made having flexural strengths (at 60–70 MPa) comparable to sintered ceramics, without the need for fibrous reinforcement or high pressure compaction. This principle is not limited to Portland cement but gives similar improvements when aluminous or sulphate cements are used. Indeed, it should apply generally to all hydraulic cements. Nor is the strengthening limited to tensile fracture; compressive strengths are elevated, being typically greater than 200 MPa.

Received 15 October; accepted 15 November 1980.

1. Soroka, I. *Portland Cement Paste and Concrete* (Macmillan, London, 1979).
2. Neville, A. M. *Properties of Concrete*, 400–403 (Pitman, London, 1963).
3. Schiller, K. K. in *Mechanical Properties of Non-Metallic Brittle Materials* (ed Walton, W. H.), 35–49 (Butterworth, London, 1958).
4. Higgins, D. D. & Bailey, J. E. *J. Mater. Sci.* **11**, 1995–2003 (1976).
5. Berger, R. L., Lawrence, F. V. & Young, J. F. *Cem. Concr. Res.* **3**, 497–508 (1973).
6. Feret, R. *Bull. Soc. Encouragement Indust. Nat.* **11** 1604 (1897).
7. Abrams, D. *Bull. Struct. Mater. Res. Lab. Lewis Inst. Chicago*, No 1 (1918).
8. Powers, T. C. *Proc. 4th Int. Congr. Chem. Cement*, Washington DC, 577–613 (US National Bureau of Standards Monograph 43, 1962).
9. Sereda, P. J., Feldman, R. F. & Ramachandran, V. S. *Proc. 7th Int. Congr. Chem. Cement*, Vol. 1, Principal Rep. V1-1 (3–44) (Editions Septima, Paris 1980).
10. Kendall, K. *Proc. R. Soc. A* **361**, 245–263 (1978).
11. Griffith, A. A. *Phil. Trans. R. Soc. A* **221**, 163 (1920).
12. Knudsen, F. P. *J. Am. ceram. Soc.* **42**, 376–387 (1959).
13. Davidge, R. W. *Mechanical Behaviour of Ceramics* (Cambridge University Press, London, 1979).
14. Brown, W. F. & Srawley, J. E. *ASTM Spec. Tech. Publ.* 410 (ASTM, Philadelphia, 1967).

15. Rootare, H. M. *Perspect. Powder Met.* **5**, 225-252 (1970).
16. De Hoff, R. T. & Rhines, R. N. *Quantitative Microscopy* (McGraw Hill, New York, 1968).
17. Winslow, D. N. & Diamond, S. J. *Mater.* **5**, 564-585 (1970).
18. Diamond, S. & Dolch, W. J. *Colloid Interface Sci.* **38**, 234-244 (1972).
19. Dullien, F. A. L. *Proc. int. Symp. Rilem/IUPAC Pore Structure and Properties of Materials* Vol 1, C173-186 (Academia, Prague, 1974).
20. European Patent Application No. 80301909.0 Imperial Chemical Industries Ltd (1980).

Conductive polymer: reticulate doping with charge-transfer complex

Jeremiasz K. Jeszka, Jacek Ulański
& Marian Kryszewski

Polymer Institute, Technical University of Łódź, 90-924 Łódź,
Żwirki 36, Poland

Additives which can form charge-transfer (CT) complexes enhance the photoconductivity of polymers. In contrast, the dark conductivity of polymers containing this type of additive increases only slightly. Several crystalline CT complexes exhibit high dark conductivity. When introduced into polymer matrix these smuld therefore significantly increase its conductivity provided that the CT complexes are ordered, because their high conductivity is related to the strong interactions between donor and acceptor stacks in the crystal. We describe here a new method of introducing microcrystalline CT complexes into polymers. It consists of growing microcrystals during film casting. The system, poly (bisphenol A carbonate)/(PC) with microcrystals of the CT complex of 7,7,8,8-tetracyanoquinodimethane (TCNQ) with tetrathiotetracene (TTT), reaches the dark conductivity of $3 \times 10^{-2} \Omega^{-1} \text{ cm}^{-1}$ at room temperature at 1 wt% of the complex. This conductivity is only slightly temperature dependent.

PC is a polymer with very low electrical conductivity but with good mechanical properties. The crystalline complex of TTT and TCNQ of stoichiometry 1:2 has a conductivity of $10^2 \Omega^{-1} \text{ cm}^{-1}$ at 300 K measured using a no-contact method at 10^{10} Hz (ref. 1). Casting of a 4 wt% solution of PC and appropriate amounts of TTT and TCNQ on *o*-dichlorobenzene on to hot (390 K) glass substrate produces high-conducting polymer films containing microcrystalline precipitations of the CT complex².

Micrographs of the film containing 1 wt% of the complex and that of its cross-section are shown in Fig. 1a and b respectively. Clearly the microcrystallites are not distributed at random but form dendrite-like paths penetrating the polymer matrix. The

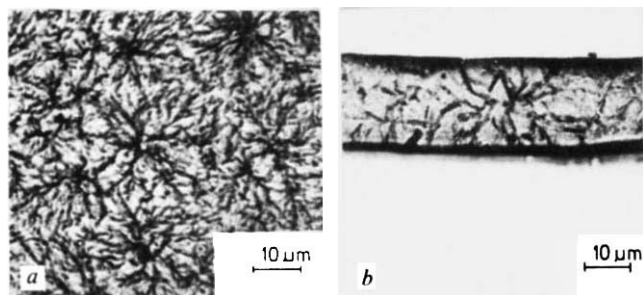


Fig. 1 Micrograph of the PC film containing 1 wt% of the TTT:TCNQ CT complex (a) and of its cross-section (b). Solutions of PC and TTT and of PC and TCNQ in *o*-dichlorobenzene were prepared separately, and after dissolving the additives at ~420K were mixed together and cast on to glass substrate at 390 K. To preserve appropriate conditions for microcrystallite growth and densification of the polymer matrix, the rate of solvent evaporation has to be controlled. PC (BDH Chemicals, laboratory reagent) was used as received, TTT and TCNQ were purified by recrystallization and the solvent was redistilled before use.

microcrystallites disappear irreversibly when a sample is annealed above 450 K, they are not observed in samples cast at temperatures higher than 450 K and in both cases no high conductivity is observed.

The dependence of the specific conductivity of the films on the content of the CT complex in the sample is shown in Fig. 2. This dependence has a yep-like characteristic and for ~1 wt% of the additive, the conductivity is increased by more than 17 orders of magnitude. When the additive content is higher, a saturation effect is observed. The microscopic investigation of these samples shows that for concentrations below 0.75 wt% the network of microcrystallites does not penetrate all the sample, while for concentrations higher than 1.5 wt% needle-like crystallites of uncomplexed TCNQ and TTT are observed.

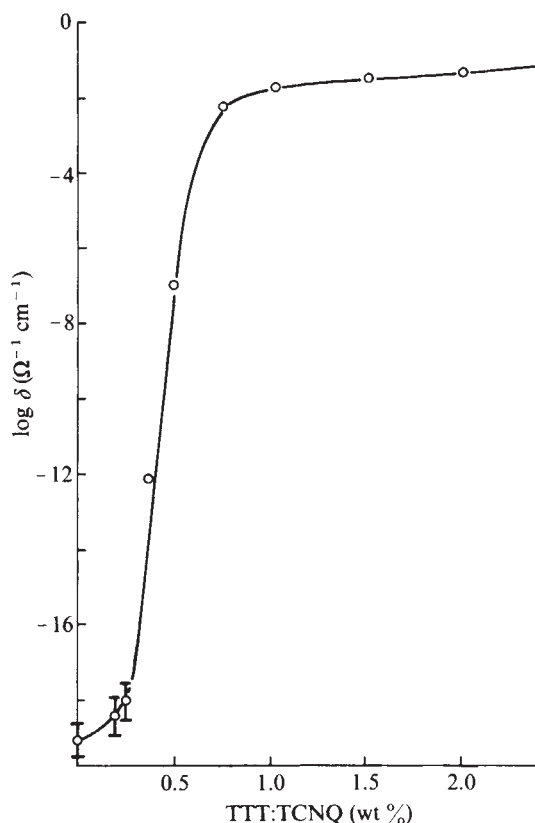


Fig. 2 The dependence of the specific conductivity of the system on the content of the TTT:TCNQ CT complex. Measurements were carried out in a vacuum chamber. For samples having high conductivity (0.5-2.0 wt% of the complex), the four-probe technique was used. Samples having higher resistivity were sandwich-type with evaporated gold electrodes and a guard ring. For samples with intermediate conductivity, both techniques give equivalent results.

The conductivity of homogeneous films obtained by casting at 450 K or by heating conducting samples above 450 K does not exceed $10^{-17} \Omega^{-1} \text{ cm}^{-1}$ even when the highest concentrations of additives are used. This indicates that the existence of the network of microcrystallites is a prerequisite for the high conductivity of the investigated material.

The temperature dependence of conductivity for samples containing 0.5, 1.0 and 2.0 wt% of TTT:TCNQ is presented in Fig. 3. Contrary to the metal-like temperature dependence of the electrical conductivity found for good conducting CT complexes of TCNQ so-called organic metals, the conductivity of the systems under study shows a weak increase with increasing temperature, diminishing with increasing content of the CT complex.

Current-voltage characteristics of highly conducting samples (concentrations above 0.75 wt%) are ohmic in the investigated temperature range up to the field strength 10^4 V m^{-1} . For higher

fields the evolved heat causes overheating of some regions which leads to destruction of the crystals and the samples becoming nonconductive. When the CT complex content is lower, the conductivity seems to be field dependent for the field strength 10^5 V m^{-1} and higher.

To find the real stoichiometry of the complex being formed in these conditions, samples with different donor-to-acceptor molar ratios were investigated. For samples containing 1 wt% of the additives, the molar ratio of TCNQ to TTT changed from 0.45 to 3.0. The conductivity measurements have shown that the highest conductivity is obtained when the molar ratio is 1.0. Also, the microscopic observations of the samples with donor-to-acceptor ratio different from one showed the presence of crystallites characteristic of an excess of uncomplexed component. This suggests that the stoichiometry of the adduct formed within the polymer is 1:1, which differs from the 1:2 stoichiometry of the naked adduct¹.

A relatively high conductivity of the samples containing 1% or more of the CT complex means that the Hall effect may be observed. The experiment was carried out at room temperature in a magnetic field of 10 T for samples containing 1% of additives. The effect was weak, but it could be determined that the excess charge carriers were electrons and their mobility could be estimated to be of the order of $10^{-1} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$.

The step-like dependence of conductivity on the amount of additive present suggests that high conductivity of this material is related to the existence of a microcrystallite network of TTT:TCNQ complex which forms conducting paths, while electrical properties of the polymeric matrix change by only a small extent. On the other hand, the fact that the temperature dependence of conductivity differs from that of highly conducting TCNQ salts indicates that there are some energetic barriers in the paths. These can arise from a separation of some crystallites by the polymer, in which case the increase of conductivity with temperature would be limited by thermal expansion of the polymeric matrix. This effect is, of course, more pronounced for smaller concentrations of the complex (see Fig. 3). It can also explain deviations from Ohm's law in current-voltage characteristics observed for complex concentrations lower than 0.5%.

The value of mobility, which is very high even compared with the mobility in the best polymer photoconductors³, can be accepted assuming the predominant role of conducting paths and taking into account high values of mobility in organic metals. Assuming one carrier per TTT:TCNQ molecule and full 1:1 complexing, one can calculate for the samples with 1 wt% of additives the density of charge carriers of the order of 10^{19} cm^{-3} and the mobility about $10^{-2} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$, while from the Hall effect we find 10^{18} cm^{-3} and $9 \times 10^{-2} \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$ respectively which, as the full complexing does not take place, seems to be a reasonable value.

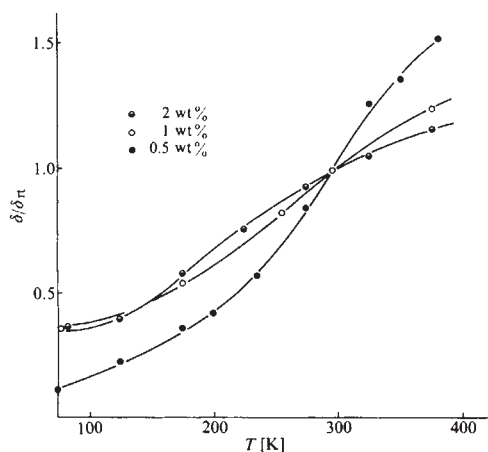


Fig. 3 The temperature dependence of electrical conductivity for the samples with different content of the CT complex. ●, 2 wt%; ○, 1 wt%; ●, 0.5 wt%

Finally, note that mechanical mixing of polymers with microcrystalline powder of conducting material is much less effective because in this case the amount of conductive additives must usually exceed 20% to achieve comparable values of conductivity⁴. The origin of this difference may be partially explained by a simple geometrical argument: for a needle of length 10 times its other dimensions, the volume is about 50 times smaller than that of a sphere of diameter equal to the needle length. If the needles are perfectly aligned, something of the order of 50 times as much weight of powder would be needed to produce the same effect as oriented needles. In our system, considerable local orientation of microcrystals is obtained by the crystallization of the complex in the polymer matrix in suitable conditions. This orientation is not perfect but a factor of ~10–20 may be expected on the basis of this argument.

The discussed mixtures, due to their high conductivity obtained at low concentration of the CT complex, combine the good application properties of commercial polymers (in contrast with polyacetylenes or $(\text{SN})_x$) with the relatively high conductivity, and form a new class of polymeric materials.

Received 7 August; accepted 14 November 1980.

1. Buravov, L. J. *Soviet Phys. JETP Lett.* **20**, 208–211 (1974).
2. Kryszewski, M., Jeszka, J. K. & Ulański, J. Polish Patent P-215189 (1979).
3. Kryszewski, M. *Semiconducting Polymers* (PWN-Polish Scientific, Warsaw, 1980).
4. Norman, R. H. *Conductive Rubbers and Plastics* (Elsevier, Amsterdam, 1970).

Interlayer formation of humin in smectites

P. Cloos, C. Badot & A. Herbillon

Groupe de Physico-Chimie Minérale et de Catalyse, UCL, and Section de Physico-Chimie Minérale, MRAC Place Croix du Sud 1, B-1348 Louvain-la-Neuve, Belgium

The stability of humin, the alkali-resistant part of humic matter in soil^{1–4}, is usually accounted for by its intimate association with inorganic soil colloids, especially swelling clays and iron compounds. Although the stability of these organo-clay complexes has been explained by interlamellar intercalation, there is evidence^{2–4} that large-sized humic acid nuclei are unlikely to penetrate layer silicates in normal soil conditions. The ability of clay minerals to promote polymerization of organic monomers has long been recognized⁵, and has been confirmed recently^{6–8}; however, the conditions in which these catalytic properties were exhibited were unlike those existing in natural soils. We now report the results of preliminary investigations suggesting that humin-like substances may form between layers of silicate clays from smaller aromatic molecules in conditions that resemble those occurring naturally.

To approach field conditions, we percolated an aqueous aniline (50 p.p.m.) solution through columns containing a 9:1 (w/w) mixture of fine sand and clay. Homoionic Fe(III), Al and Ca exchanged forms of raw montmorillonite from Camp-Berteau (Morocco) were used. An appropriate percolation velocity and clear effluents were obtained by alternating in the columns (height, 140 mm; diameter, 45 mm) ~1-cm thick layers of the sand-montmorillonite mixture and of neat sand, deposited on a 3–4 cm-thick bottom layer of sand supported by glass wool. Apart from interruptions at the weekends, a 50-ml portion of the dilute aniline solution was daily percolated through each column.

During the first treatment, similarly to observations⁹ made on 1% clay suspensions, light-green trails and spots developed in the top Fe(III)-clay/sand layers. They turned progressively to green-blue, then to dark greenish-grey whilst increasing in size and extending slowly towards the bottom of the column as more aniline was added. The Al and Ca saturated samples did not

Table 1 C and N contents and/or basal spacings at increasing dehydration levels

Sample	C (%)	N (%)	C/N	Basal spacings $d(001)$ (Å)		
				Air-dry 20 °C	Degassed (2h) 20 °C	in vacuo 80 °C
Fe(III)-montmorillonite	—	—	—	15.5	11.8	11.3
Aniline-Fe(III)-montmor	1.12	0.22	6/1	14.2	12.1	11.9
Fe(III)-montmorillonite*	—	—	—	16.6	11.0	10.6
Aniline-Fe(III)-montmor*	1.05	0.37	3.3/1	16.3	12.5	11.3
Fe(III)-montmorillonite†	—	—	—	13.2	10.3	10.3
Aniline-Fe(III)-montmor†	1.13	0.18	7.3/1	13.4	10.65	10.65
DMSO extract‡	38.69	4.74	9.5/1	—	—	—

* Treated with a 0.1 M solution of HCl.

† Treated with a 0.1 M solution of $\text{Na}_4\text{P}_2\text{O}_7$.

‡ After HF attack of the aniline-Fe(III)-montmorillonite complex.

show any change in colour; after a few percolations, the aniline concentration of the effluents equalled that of the treating solution, and washing with distilled water removed most of the organic added. In contrast, no aniline could be detected in the effluents of the column containing Fe(III)-clay, the bottom layer of which had not yet been reached by the green-grey zones at the end of six weeks, when the experiment finished.

After air-drying the column at room temperature, the dark-coloured material was collected and the colloidal organo-clay fraction was separated by sedimentation and freeze-dried. Treating it with a 0.1 M solution of HCl or a 0.1 M alkaline solution of $\text{Na}_4\text{P}_2\text{O}_7$ did not modify either its colour (7.5 GY 4/1 to 10 GY 4/1 on the soil colour chart) or its carbon content (Table 1). Hence, according to the conventional solubility criteria¹⁻⁴, the dark-coloured organic substance firmly held by the clay may be termed humin-like material.

Several attempts to extract this humin-like material with various organic solvents failed. Only with dimethyl sulphoxide (DMSO) could we extract part of it from a sample previously attacked with a dilute HF solution. The dark grey to black powder, left after evaporation of the solvent, still had an ash content of 23% for a carbon content of nearly 40% (Table 1).

Assuming interlamellar intercalation, the amount of aniline (Table 1) fixed as a monomolecular layer would cover less than 10% of the internal clay surface⁸. Nevertheless, the basal spacings (Table 1) determined on oriented clay films in a vacuum camera, are systematically higher for the increasingly dehydrated organo-clay complex than for the corresponding neat clay.

A stronger argument for interlamellar intercalation is provided by the IR and ESR spectra (Fig. 1) of the aniline-Fe(III)-montmorillonite complex. The IR spectra, recorded on self-supporting clay films, show evidence of hydration water (ν_{OH} near 3,500, δ_{OH} near 1,640 cm^{-1}), and are characterized by broadening of the aromatic ring vibrations in the 1,600–1,500 cm^{-1} range, by the development of bands near 1,320 and 1,255 cm^{-1} and of a broad charge-transfer absorption extending from roughly 2,000 to beyond 4,000 cm^{-1} . The ESR spectra exhibit a sharp radical signal near the position of a free spinning electron (reference), superimposed on the broad Fe(III) signal. These features are typical for the so-called Type II complexes¹⁰⁻¹³, which reportedly form exclusively in the interlayer space of smectites¹⁴. Their formation generally occurs in dehydrating conditions, through a π interaction between certain aromatic compounds and some transition metal cations occupying exchange sites. This interaction leads to electron transfer resulting in the reduction of the metallic cation and the formation of aromatic radical cations able to participate in polymerization reactions⁶⁻⁸.

The typical bands of the aniline-Fe(III)-montmorillonite Type II complex persist in the IR spectrum of the DMSO extract (Fig. 1). An additional absorption near 1,665 cm^{-1} , compatible with quinonoid C=O stretching, suggests further oxidation, as do the enhancement of the radical signal (Fig. 1) and the relative decrease in nitrogen content (Table 1) on treating the organo-clay complex with $\text{Na}_4\text{P}_2\text{O}_7$ or DMSO. Similarly, the relative

increase in nitrogen content of the acid-treated sample is explainable by partial oxidative hydrolysis, the protonated amino moiety being retained on the clay by exchange with Fe(III).

Ferric iron was present mostly as hydroxycations ($\text{Fe}(\text{OH})^{2+}$, $\text{Fe}(\text{OH})_2^+$) or polymeric species of these, as indicated by the apparent cation exchange capacity of 184 mEq. per 100 g, instead of 100 mEq. per 100 g when determined on NH_4 -montmorillonite at pH 7. Ageing for more than 10 months, in aqueous suspension or in ambient conditions, did not prevent Fe(III)-montmorillonite from interacting with aniline as described above. Moreover, aniline seemed to be the first aromatic compound to form a Type II complex on a smectite in aqueous medium, evolving to polymeric species in near-pedogenic conditions.

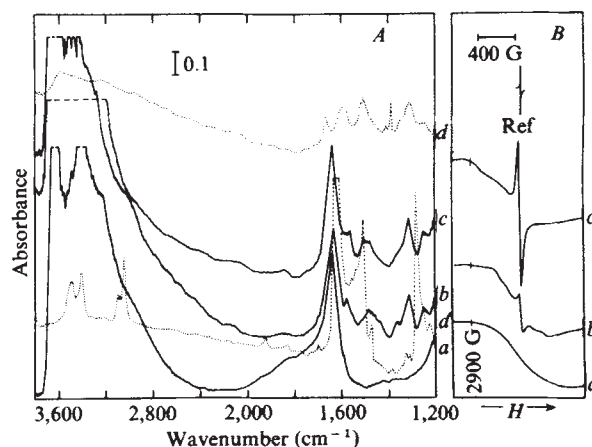


Fig. 1 IR (A) and electron spin resonance (B) spectra for: a, neat Fe(III)-montmorillonite; b, aniline-Fe(III)-montmorillonite complex; c, same after $\text{Na}_4\text{P}_2\text{O}_7$ treatment. Spectra a' and d refer to aniline (in CCl_4 solution) and to the DMSO extract of the aniline-Fe(III)-montmorillonite complex (in KBr pellet), respectively.

The above observations demonstrate that interlayer humin formation in smectitic soils or sediments can take place through Type II complexes at the expense of small-sized aromatic molecules. Although aniline is not very representative of biologically generated, natural aromatics—much less than phenolic compounds^{1,3}, for instance—its behaviour is probably not unique. The important result is the interlamellar polymerization of a simple aromatic monomer, at low concentration in aqueous solution and at low temperature on a smectite containing Fe(III) exchange forms. This type of reaction might well be an alternative for the intercalation of large-sized, polymeric humic acid nuclei, highly improbable in natural conditions.

Received 30 September; accepted 18 November 1980.

1. Schnitzer, M. & Khan, S. V. *Humic Substances in the Environment* (Marcel Dekker, New York, 1972).
2. Duchaufour, P. *Pédologie* Vol. 1 (Masson, Paris, 1977).
3. Hayes, M. H. B. & Swift, R. S. in *The Chemistry of Soil Constituents* Ch. 3 (eds Greenland, D. J. & Hayes, M. H. B.) (Wiley, New York, 1978).

4. Theng, B. K. G. *Formation and Properties of Clay-Polymer Complexes* (Elsevier, Amsterdam, 1979).
5. Theng, B. K. G. *The Chemistry of Clay Organic Reactions* (Hilger, London, 1974).
6. Mortland, M. M. & Halloran, L. J. *Am. J. Soil Sci. Soc.* **40**, 367 (1976).
7. Stoessel, F., Guth, J. L. & Wey, R. *Clay Miner.* **12**, 255 (1977).
8. Cloos, P., Moreale, A., Broers, C. & Badot, C. *Clay Miner.* **14**, 307 (1979).
9. Moreale, A. & Van Bladel, R. *Clay Miner.* **14**, 1 (1979).
10. Mortland, M. M. & Pinnavaia, T. J. *Nature phys. Sci.* **229**, 75 (1971).
11. Cloos, P., Vande Poel, D. & Camerlynck, J. P. *Nature phys. Sci.* **243**, 54 (1973).
12. Pinnavaia, T. J., Hall, P. A., Cady, S. S. & Mortland, M. M. *J. phys. Chem.* **78**, 994 (1974).
13. Rupert, P. J. *J. phys. Chem.* **77**, 784 (1973).
14. Doner, H. E. & Mortland, M. M. *Science* **166**, 1406 (1969).

Compositional convection and stratification of Earth's core

David R. Fearn* & David E. Loper

Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, Florida 32306

Traditionally, studies^{1,2} of the possible stratification of the outer core of the Earth have been based on thermal buoyancy, with the effects of composition being implicitly ignored. However, recent work^{3,4} has demonstrated the importance of compositional buoyancy in providing the power necessary to sustain the geodynamo. It seems likely, therefore, that the effects of composition are also important in determining the stratification of the core. We present here a study of the relative importance of thermal and compositional buoyancy in the core. We briefly review the idea that the growth of an iron-rich inner core leads to the release of material which is depleted in iron and therefore buoyant. We then calculate the density profile which results if both this light material and the latent heat released as the inner core grows are redistributed by diffusion. It is found that not only are compositional effects important, they dominate thermal effects in determining the stratification of the core. Finally, using parameter estimates relevant to the core, we show that the diffusive density profile is unstable throughout the outer core except for a layer some 70 km deep adjacent to the mantle-core boundary. We therefore expect that convection fills the entire outer core with the possible exception of a thin layer below the mantle-core boundary.

It has long been believed that the Earth is in thermal equilibrium, with the internal generation of heat by radioactive decay balancing the surface heat flux. Recently, Stacey⁵ pointed out that such an equilibrium cannot be maintained because the strength of the radioactivity decays with time and he argued that a significant portion of the surface heat flux is due to the gradual cooling of the entire Earth. This idea is strongly supported by studies⁶⁻⁸ using parameterized models of thermal convection.

If the Earth is cooling, there will come a time when the core fluid reaches its freezing temperature and solid begins to freeze out. If the freezing temperature gradient is steeper than the actual temperature gradient, then freezing will first occur at the centre of the core, and the frozen solid will accumulate to form a solid inner core⁹. The freezing process results in the release of light material as well as of latent heat. Both are sources of buoyancy and have been proposed^{10,11} as the cause of the convective motions necessary for dynamo action.

The core is not a pure material but is composed mainly of iron together with a significant fraction (8–20%) of some lighter material¹². The precise nature of this material is not important for our purposes and it is sufficient to model the core as a binary alloy of a heavy metal and a light non-metal. We shall assume the mass fraction ξ of light material to be less than that required for an eutectic mixture ($\xi < \xi_e$) as the case $\xi > \xi_e$ proposed by Braginsky¹¹ has been shown to be geophysically implausible¹³. Metallurgical studies¹⁴ have shown that the solid which freezes out when $\xi < \xi_e$ contains more of the heavy component and is therefore more dense than the surrounding liquid. By conser-

vation of light material, the freezing process must, therefore, leave a light, compositionally buoyant residue in the liquid adjacent to the freezing interface. Thus, as the inner core grows, light material is released and it is this compositional buoyancy which Braginsky¹¹ proposed as the source of the energy for the geodynamo.

The energetics of Braginsky's proposal have been calculated^{3,4} by estimating the rate of release of gravitational potential energy resulting from the net motion of light material upward and heavy material downward, assuming the inner core to have grown at a constant rate over the lifetime of the Earth. This mechanism was found to be capable of supporting a large toroidal magnetic field through dynamo action. Here we study the dynamics of Braginsky's proposal by estimating the relative strengths of the buoyancy forces produced by temperature and compositional differences within the core.

If the inner core is growing at a mass rate $\dot{M}$ and the change in composition on freezing is denoted by ξ_i , Loper and Roberts¹⁵ have shown that the upward flux I of the light material and the thermal diffusive flux Q of heat into the outer core are given by

$$I = \xi_i \dot{M} \quad \text{and} \quad Q = L \dot{M}, \quad (1)$$

where L is the latent heat. The light material must be redistributed throughout the fluid outer core (causing a gradual increase in the mass fraction of light material) and the heat must be transported up to the mantle-core boundary. There are two possible mechanisms for redistribution: diffusion and convection. The mechanism which occurs is determined by the density gradient. A necessary condition for the diffusive state to be unstable to a convective instability is that a parcel of fluid displaced upward isentropically and without change of composition be less dense than its surroundings;

$$(\partial \rho / \partial r)_{s, \xi} < d\rho_o / dr \quad (2)$$

where r is the radial coordinate and a subscript o denotes the static diffusive state. Loper and Roberts¹⁵ have shown that for the large length scales in the core, the density difference required to overcome diffusive processes associated with convective motions is very small and to a good approximation equation (2) is also sufficient for convective instability.

Let us adopt equations of state for density and entropy of the form

$$d\rho = \beta \rho d\rho - \alpha \rho dT - \bar{\delta} \rho^2 d\xi$$

and

$$dS = -(\alpha/\rho) d\rho + (C_p/T) dT + \bar{s} d\xi$$

(see ref. 16). We may now rewrite equation (2) as

$$\alpha \rho dT_o / dr + \bar{\delta} \rho^2 d\xi_o / dr < (\alpha^2 T / C_p) d\rho / dr \quad (3)$$

where C_p is the specific heat capacity at constant pressure, α is the thermal expansion coefficient, and $\bar{\delta}$ is the specific volume change with composition.

In the diffusive state the radial fluxes Q_o of heat and I_o of light material are related to the diffusive gradients by

$$I_o = -\rho D A [d\xi_o / dr + (\bar{\delta} / \bar{\mu}) d\rho / dr] \quad (4)$$

and

$$Q_o = -k A dT_o / dr \quad (5)$$

The constitutive laws (4) and (5) are equivalent to equations (58.11) and (58.12) of Landau and Lifshitz¹⁷ with $Q_o = A \hat{\mathbf{e}} \cdot [\mathbf{q} - (\bar{\mu} + \bar{s} T) \mathbf{i}]$ and the coefficient k_T set equal to zero. Here $A = 4\pi r^2$ is the area of the sphere of radius r , k is the thermal conductivity, D is the material diffusion coefficient and $\bar{\mu}$ is the energy of mixing. At the inner-core boundary I_o and Q_o are given by equation (1), whereas at the mantle-core boundary we assume there is no flux of light material ($I_o = 0$), and a specified heat flux Q_m . The flux Q_m results from the diffusion of the latent heat $L \dot{M}$ across the outer core, plus the heat capacity released by the cooling of the core and any heat generated by radioactive decay, ohmic heating, and so on. If we assume the internal heat sources to be uniformly distributed throughout the outer core,

* Present address: Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, UK.

then

$$Q_o = [(r_m^3 - r_i^3)\dot{M} + (r^3 - r_i^3)Q_m]/(r_m^3 - r_i^3) \quad (6)$$

where r_m is the radius of the outer core and r_i is the radius of the inner core. Similarly, assuming the material flux I_o distributes the light material uniformly throughout the outer core, we have

$$I_o = \xi_i \dot{M} (r_m^3 - r^3)/(r_m^3 - r_i^3). \quad (7)$$

Combining equations (3)–(7) and assuming a hydrostatic balance, $dp/dr = -\rho g$ we obtain

$$\frac{\alpha^2 T}{C_p} + \frac{(\delta\rho)^2}{\bar{\mu}} < \frac{\left[(r_m^3 - r^3)\dot{M} \left(\frac{\alpha L}{k} + \frac{\delta\xi_i}{D} \right) + (r^3 - r_i^3) \frac{\alpha}{k} Q_m \right]}{4\pi r^2 (r_m^3 - r_i^3) g} \quad (8)$$

The terms on the left-hand side of equation (8) represent the stabilizing effects of the conduction of heat down the adiabat and the diffusion of material induced by the pressure gradient (barodiffusion), respectively, while the terms on the right-hand side represent the destabilizing effects of the release of latent heat and of light material associated with the freezing process and the heat flux out through the mantle–core boundary. This is a generalization of equation (14) of Loper and Roberts¹⁸.

The instability criterion (8) can be considerably simplified for parameter values thought to be relevant to the core. Loper and Roberts¹⁸ have shown that $\alpha^2 T/C_p \ll (\delta\rho)^2/\bar{\mu}$ and $\alpha L/k \ll \delta\xi_i/D$ at the inner-core boundary. Using the same sources as Loper and Roberts¹⁸ for the parameter estimates, it can be shown that these inequalities hold good throughout the outer core. We can also obtain a measure of the importance of the term involving Q_m by comparing its maximum value (at $r = r_m$) with the left-hand side of equation (8). Using parameter estimates appropriate at the mantle–core boundary, with $Q_m = 5 \times 10^{12}$ W, we find that $[\alpha Q_m/gk4\pi r_m^2]/[(\delta\rho)^2/\bar{\mu}] \approx 6 \times 10^{-2}$. Thus we may conclude that thermal effects are not of primary importance in determining the stability of the core. To a first approximation, the thermal terms may be neglected and equation (8) reduces to

$$\frac{(\delta\rho)^2}{\bar{\mu}} < \frac{(r_m^3 - r^3)\delta\xi_i \dot{M}}{4\pi r^2 (r_m^3 - r_i^3) g D}. \quad (9)$$

This is the condition for the diffusive state to be unstable to convective overturning. It depends only on compositional effects and is equivalent to the condition $d\xi_o/dr < 0$.

The diffusive flux of light material has two components (see equation (4)). The first is the familiar flux down the compositional gradient which acts to make the composition homogeneous (minimizing chemical potential energy). If this were the only component, instability would occur throughout the core for any positive growth rate $\dot{M}$. The second component is the barodiffusive flux which acts to separate out the light component from the heavy component (minimizing gravitational potential energy). Together the two fluxes act to redistribute light material in such a way as to minimize the total (gravitational plus chemical) potential energy. If $\dot{M} = 0$ and the net flux of light material in the core is zero, we can see from equation (4) that the compositional gradient $d\xi_o/dr = -(\delta/\bar{\mu}) dp/dr > 0$ is stable. The mass rate of growth $\dot{M}$ must exceed some critical value before the compositional gradient changes sign. As $\dot{M}$ is increased we can see from equation (9) that this happens first at the inner-core boundary ($r = r_i$). As $\dot{M}$ is further increased a layer $r_i \leq r \leq r_t < r_m$ develops in which $d\xi_o/dr < 0$. Within this layer, the diffusive solution is unstable and convection occurs.

Loper and Roberts¹⁸ have shown that at $r = r_i$, inequality (9) is satisfied by a wide margin, implying convection in the lower part of the outer core. On the other hand, at $r = r_m$ the right-hand side of equation (9) vanishes so the diffusive density profile must be stable in a layer adjacent to the mantle–core boundary. We determine the depth $h = r_m - r_t$ of this stable layer by finding that value r_t of the radius for which the left- and right-hand sides of equation (9) are equal. Solving for h we find

$$h = \rho^2 \delta D g V / \xi_i \dot{M} \bar{\mu}, \quad (10)$$

where $V = 4\pi(r_m^3 - r_i^3)/3$ is the volume of the outer core and we

have assumed $h \ll r_m$. Using parameter estimates relevant to the vicinity of the mantle–core boundary (with $D \approx 6 \times 10^{-9} \text{ m}^2 \text{ s}^{-1}$) we find the depth of the stably stratified layer to be ~ 70 km. The uncertainties in the estimates of the various parameters (particularly D and $\bar{\mu}$) mean that this figure should be regarded as only an order-of-magnitude estimate of the depth of the stable layer.

A simple calculation equating the kinetic energy of a convecting fluid parcel to the potential energy gained as this parcel penetrates the stable layer suggests that convection will penetrate only a few metres into the stable layer if it is already established. However, if the stable layer were forcibly mixed by some cataclysmic process (such as core formation), the combined action of barodiffusion and the normal diffusive process would act to re-establish the layer, but only very slowly. In fact, it would take a length of time equal to the age of the Earth to establish a layer several tens of kilometres thick by these processes. Further, this process might be severely inhibited if convective motions disrupt the newly forming stable layer before a strongly stabilizing gradient can be established. In view of these caveats, the recent claim¹⁹ that the top of the outer core is stably stratified must be treated with caution.

We thank Paul Roberts for his helpful discussions. This work was supported by the NSF under grant no. EAR78-00818 and is contribution No. 160 of the Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, Florida.

Received 16 September; accepted 27 November 1980.

1. Stacey, F. D. *Physics of the Earth* 2nd edn, Ch. 7 (Wiley, New York, 1977).
2. Jacobs, J. A. *The Earth's Core* Ch. 3 (Academic, London, 1975).
3. Gubbins, D. J. *Geophys.* **43**, 453–464 (1977).
4. Loper, D. E. *Geophys. J. R. astr. Soc.* **54**, 389–404 (1978).
5. Stacey, F. D. *Phys. Earth planet. Inter.* **22**, 89–96 (1980).
6. Sharpe, H. N. & Peltier, W. R. *Geophys. J. R. astr. Soc.* **59**, 171–205 (1979).
7. Schubert, G., Stevenson, D. & Cassen, P. J. *geophys. Res.* **85**, 2531–2538 (1980).
8. Davies, G. F. J. *geophys. Res.* **85**, 2517–2530 (1980).
9. Jacobs, J. A. *Nature* **172**, 297 (1953).
10. Verhoogen, J. *Geophys. J. R. astr. Soc.* **4**, 276–281 (1961).
11. Braginsky, S. I. *Doklady Akad. Nauk SSSR* **149**, 8–10 (1963).
12. Brett, R. *Rev. Geophys. Space Phys.* **14**, 375–383 (1976).
13. Fearn, D. R. & Loper, D. E. *Geophys. Astrophys. Fluid Dyn.* (submitted).
14. Chalmers, B. *Principles of Solidification* (Wiley, New York, 1964).
15. Loper, D. E. & Roberts, P. H. *Geophys. Astrophys. Fluid Dyn.* **16**, 83–127 (1980).
16. Loper, D. E. & Roberts, P. H. *Geophys. Astrophys. Fluid Dyn.* **9**, 289–321 (1978).
17. Landau, L. D. & Lifshitz, E. M. *Fluid Mechanics* (Pergamon, Oxford, 1959).
18. Loper, D. E. & Roberts, P. H. *Phys. Earth planet. Inter.* (in the press).
19. Whaler, K. A. *Nature* **287**, 528–530 (1980).

Significance of contrasting magmatism in North East Africa and Saudi Arabia

N. B. W. Harris & I. G. Gass

Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK

The Red Sea is a north–south rift that split the Arabian–Nubian shield into eastern (Arabian) and western (Egypt and Sudan) segments by seafloor spreading less than 30 Myr ago¹. Because geological studies in the region have generally been confined by national frontiers, a comparative study of magmatic evolution on either side of the Red Sea has not, so far as we are aware, previously been undertaken, and current models for Arabian–Nubian shield evolution suggest broadly similar histories for the basement on both sides of the Red Sea. However, although the igneous rock geochronology on both sides of the Red Sea indicate that Precambrian magmatic histories were broadly similar, since then they have been quite distinct in that the African crust has been invaded episodically by quartz syenites, nepheline syenites and peralkaline granites between 500 and 80 Myr; no such igneous activity is evident in Arabia (Fig. 1). We propose here that the Red Sea line has had structural significance from at least the end of the Cambrian though whether it defines a line of contrasting upper mantle or lower crust geochemistry, or of contrasting crustal rigidity cannot yet be established.

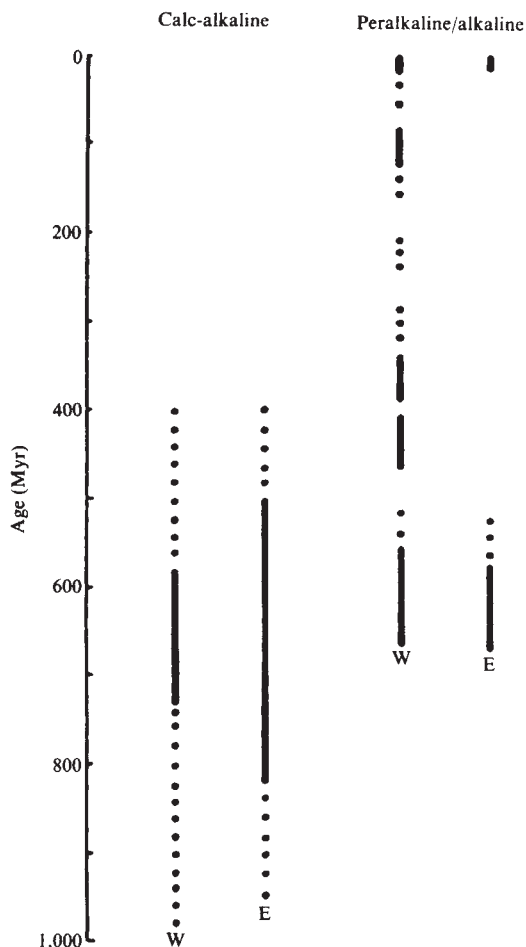


Fig. 1 Age spans of calc-alkaline and peralkaline/alkaline magmatism in the Arabian-Nubian shield. Solid lines from U-Pb zircon or Rb/Sr isochron ($N > 3$) data with $2\sigma < 10\%$. Dotted lines from K/Ar or other Rb/Sr data. E, Eastern shield (Saudi Arabia and Sinai). W, Western shield (Egypt and Sudan). All ages have been recalculated to IUGS recommended constants².

Nearly 600 age determinations have been published from the Arabian-Nubian shield, about two-thirds of which are from Saudi Arabia. Although most of these are isolated K/Ar or Rb/Sr ages, and cannot be considered reliable indicators of geological events, available U-Pb zircon and Rb/Sr isochron ages with low errors indicate that on both sides of the Red Sea calc-alkaline igneous activity was more or less continuous between 840 and 500 Myr, being initially dioritic or tonalitic in character until about 700 Myr when granitic magmatism became predominant. There is evidence of sporadic magmatism as early as 980 Myr, and a single three-point isochron of Saudi Arabian metabasalt at 1,165 Myr (ref. 3). Due to daughter isotope leakage, the age of cessation of calc-alkaline magmatism is difficult to define, and some activity may have occurred between 500 and 400 Myr on either side of the Red Sea. However, in contrast to calc-alkaline magmatism, peralkaline and alkaline magmatism have differing age distributions in Africa and Arabia (Fig. 2). In north-west Arabia, excluding Tertiary magmatism associated with Red Sea rifting, peralkaline magmatism seems to be restricted to a geologically brief period³⁻⁶ between 640 and 540 Myr whereas in north-east Africa alkaline intrusions were not only emplaced between 600 and 540 Myr (refs 7, 8), but continued to be so at irregular intervals until 80 Myr ago⁸⁻¹⁰. Many of the available dates for this magmatism are from the K/Ar method, but eight have been confirmed by Rb/Sr isochrons and six more by Rb/Sr mineral ages.

Although the data bank of reliable ages of alkaline intrusions Arabia is not as large as in north-east Africa, there are two

independent indications that magmatism did not occur in the eastern shield between 540 and 30 Myr.

(1) In Egypt and north-east Sudan, the mainly Mesozoic Nubian sandstone is intruded by alkaline ring complexes^{11,12}, whereas peralkaline intrusions in Saudi Arabia are not known to penetrate Phanerozoic sedimentary formations.

(2) In the western shield there is a critical change in magma composition at ~500 Myr. Pre-500 Myr peralkaline granites with smaller volumes of quartz syenites and peralkaline granites were emplaced^{7,9,13}. After 500 Myr the dominant rock-type is nepheline syenite with smaller volumes of quartz syenite and peralkaline granite^{9,13,14}. This depletion in silica of the younger intrusives indicates a change of either depth or volatile content in the source environment. Analyses from 17 separate alkaline intrusions from Saudi Arabia^{15,16} indicate peralkaline granites which are strongly quartz normative which therefore correlate with the pre-50 Myr group from Saudi Arabia. Moreover nepheline-bearing plutons have not been recorded in Arabia so far as we are aware. The absence of silica undersaturated magmas from Saudi Arabia supports the conclusion that the post-500 Myr group from Egypt and Sudan have no equivalent in Saudi Arabia.

Therefore, although there was little if any magmatic activity in Saudi Arabia and Sinai between 500 and 30 Myr, in Egypt and north-east Sudan there were spasmodic and seemingly random high level injections of peralkaline and alkaline magmas throughout this period.

Geochemical considerations of peralkaline granites in general and of Arabian peralkaline granites in particular indicate that such magmas originate by partial melting at depth in the presence of non-hydrous volatiles^{16,17}. Of the 16 available Rb/Sr isochrons of peralkaline and alkaline granites from both sides of the shield, 12 indicate initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios between 0.701 and 0.704 implying a source in the upper mantle or lower crust. Four remaining isochrons show ratios between 0.705 and 0.710 which suggest either crustal contamination or a source of high Rb/Sr ratio.

However, as most, if not all, peralkaline magmatism from the shield originates from either the lower crust or upper mantle, its absence in the Arabian segment after 500 Myr may be attributed to: the westward migration of an upper mantle thermal energy source; a more rigid crust to the Arabian segment; or variation in upper mantle geochemistry. These alternatives are now briefly evaluated.

If there has been a westward migration of a high heat-flow region in the upper mantle it could result in an east-west age trend of peralkaline intrusions within the African crust. No such trend is observed and intrusions with age differences of 300–500 Myr are found in juxtaposition^{9,13}.

Alternatively, it is possible that the Arabian crust became impervious to magmas. Although magma can rise through most of the crust by stokesian diapiric movement, its emplacement into the upper crust is greatly facilitated by the presence of active fractures. It is relevant therefore that fault trends differ in the two segments. In Egypt and the Sudan, an EW–ENE trend is most obvious and controls the emplacement of alkaline intrusions¹³, whereas in Arabia the dominant NW–SE Nadj system¹⁸ was active up to 500 Myr ago and is associated with the emplacement of peralkaline granites in Saudi Arabia^{15,16}. So, as both segments are cut by faults which were active during the periods of magmatic emplacement, crustal imperviousness to magmatic penetration seems an unlikely explanation for the lack of Phanerozoic peralkaline magmatism in Arabia.

The formation of silica undersaturated alkaline intrusions in the western shield implies the existence of a deep source region low in H_2O and possibly rich in CO_2 (ref. 19). The absence of similar intrusions in the eastern shield may indicate a distinct upper mantle geochemistry and, if this proves to be the case, the distinction existed 450 Myr before the opening of the Red Sea. Although it is possible that the boundary between two upper mantle regions of distinct geochemistry coincidentally become the site of lithospheric rupture, the following model seems to

provide a more obvious link between upper mantle heterogeneity and subsequent rifting.

There is strong and increasing evidence^{20,21} that the continental crust of the Arabian-Nubian shield was produced by the evolution and accretion of originally intra-oceanic island arc systems during the Upper Proterozoic. This process was complete by about 500 Myr at which time the region attained a continental character, but this was obviously reached in some places before others¹⁶. The presence of peralkaline granites on both sides of the shield before 500 Myr indicates that the upper mantle or lower crust was able to provide magmas which were emplaced as peralkaline granite melts. Structural and geochemical evidence suggests the Arabian shield was underthrust by one or more Upper Proterozoic, easterly dipping subduction zones^{3,21}. In this event, a mantle wedge beneath Arabia and above the subduction zones would be hydrated due to seawater being expelled from the subducted oceanic plate which in time would lower the liquidus temperature of the sub-Arabian peridotite mantle, and the melts produced would be enriched in H₂O

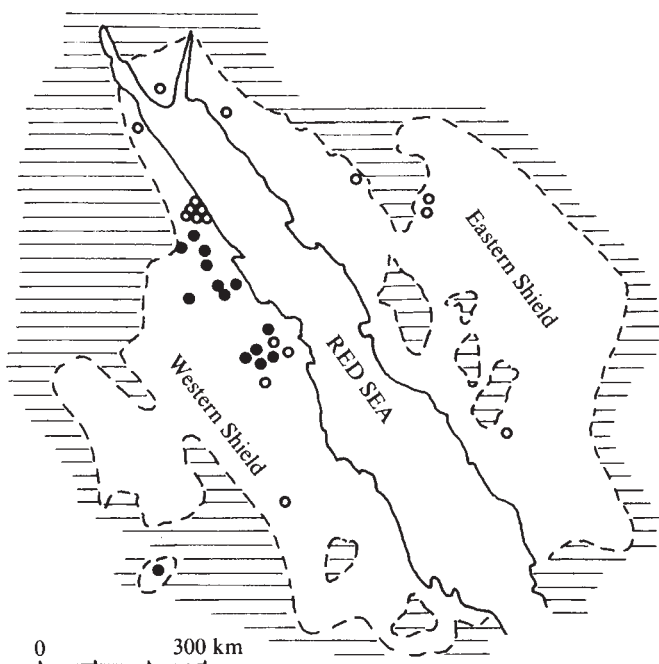


Fig. 2 Distribution of dated samples from alkaline and peralkaline intrusions. ○, 660–540 Myr. ●, 480–80 Myr. Horizontal shading marks Phanerozoic cover.

and therefore be of calc-alkaline affinity. Peralkaline magmatism, which requires conditions of low P_{H_2O} , would only be initiated after subduction had ceased. However, the sub-Arabian mantle and lower crust would be depleted in components contributing to low temperature melts, and therefore could not provide peralkaline melts after ~500 Myr until the opening of the Red Sea. In contrast, the sub-African mantle, which had not been extensively depleted by such late stage calc-alkaline magmatism retained its low melting point components and continued to produce small quantities of alkaline magma throughout the Phanerozoic. During the Tertiary the Red Sea rift developed along the line separating depleted and non-depleted upper mantle, which in turn may coincide with a Proterozoic suture zone.

Received 2 June; accepted 28 November 1980.

1. Gass, I. G. *Nature* **165**, 722–724 (1977).
2. Steigler, R. H. & Jager, E. *Earth planet. Sci. Lett.* **36**, 359–363 (1977).
3. Fleck, R. J., Greenwood, W. R., Hadley, D. G., Anderson, R. E. & Schmidt, D. G. *U.S. geol. Surv. Saudi Arabian Proj. Rep. No. 245* (1979).
4. Aldrich, L. T., Brown, G. F., Hedge, C. & Marvin, R. F. *U.S. geol. Surv. Proj. Rep. No. 240* (1978).

5. Baubron, J. C., Delfour, J. & Viallette, Y. *Open-file Rep. 76-JED-22* (Bureau de Recherches Geol. Miniers, 1976).
6. Bielski, M., Jager, E. & Steinitz, G. *Contr. Miner. Petrol.* **70**, 159–165.
7. Roger, J. J. W., Ghuma, M. A., Nagy, R. M., Greenberg, J. K. & Fullager, P. D. *Earth planet. Sci. Lett.* **39**, 109–117 (1978).
8. Lutz, T. M., Foland, K. A., Faul, H., Friedman, I. & Serencsits, C. McC. *U.S. geol. Surv. Open-file Rep. 78-701*, 465–467 (1978).
9. Cavanagh, B. J. thesis, Univ. Leeds (1979).
10. Hashad, A. H. *Bull. Inst. appl. Geol. Jeddah*, **3**, 31–46 (1980).
11. Vail, J. R. *Geol. J.* **B20**, 97–114 (1976).
12. Gass, I. G. thesis, Univ. Leeds (1955).
13. Serencsits, C. McC., El Ramly, M. F., Faul, H., Foland, K. A. & Hussein, A. A. *Annls. geol. Surv. Egypt* (in the press).
14. El Ramly, M. F., Budanov, V. I., Armanious, L. K. & Devenui, N. E. *United Arab Repub. Geol. Surv. Pap.* 52 (1969).
15. Stoesser, D. B. & Elliot, J. E. *U.S. geol. Surv. Proj. Rep. No. 265* (1979).
16. Harris, N. B. W. & Marriner, G. F. *Lithos* **13**, 325–337 (1980).
17. Bailey, D. K. in *The Alkaline Rocks* (ed. Sovensen, H.) (Wiley, New York, 1974).
18. Moore, J. M. J. *geol. Soc. Lond.* **136**, 441–454 (1979).
19. Wyllie, P. J. & Huang, W. L. *Contr. Miner. Petrol.* **54**, 79–107 (1976).
20. Gass, I. G. *geol. Soc. Lond.* **136**, 129–138 (1977).
21. Greenwood, W. R., Hadley, D. G., Anderson, R. E., Fleck, R. J. & Schmidt, D. L. *Phil. Trans. R. Soc. A280*, 517–526 (1976).

Solar energy conversion by chloroplast photoelectrochemical cells

Ravindra Bhardwaj, Rong L. Pan
& Elizabeth L. Gross

Department of Biochemistry, The Ohio State University,
484 West 12th Avenue, Columbus, Ohio 43210

Solar energy can be captured by photochemical or photoelectric processes in which a gradient of chemical potential can be generated using photochemical redox reactions. Photosynthesis is the most efficient system for quantum conversion and storage of solar energy¹. Various biomimetic systems have been constructed as models for the conversion of light to electrical energy^{2–11}. These biological photoelectrochemical cells have utilized microcrystalline chlorophyll *a* (Chl *a*)², chlorophyll liquid crystals deposited on the electrodes^{3–5} or incorporated into filters⁶, bilayer lipid membranes⁷, chloroplast membranes⁸, chloroplasts⁹ and bacterial reaction centres^{10,11}. The power conversion efficiencies are mostly very low ($\approx 0.002\%$). We describe here a chloroplast photoelectrochemical cell which generates large photovoltage and photocurrent, and has a power conversion efficiency close to 1%.

We have previously developed a solar cell using photosystem I (PSI) particles¹². Our rationale was that charge separation occurs within the reaction centre of PSI particles and produces a

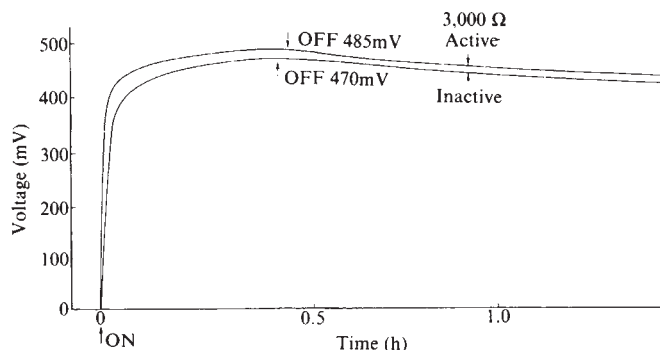


Fig. 1 Time course of potential development. Chloroplasts containing 80 μg Chl were aspirated onto a 1-cm² area (1.2- μm pore diameter) Metrical filter. The cell contained 80 mM Tricine (pH 7.0) + 100 mM NaCl + either 3.0 mM FMN or 0.5 mM DCIP. The FMN side (front surface area = 8 cm²) was illuminated with white light ($I = 4,600 \text{ W m}^{-2}$). The potential was measured across a 3,000 Ω load resistance.

Table 1 The effect of buffers on the chloroplast solar cell

Buffer	Buffer concentration (mM)	Condition	DCIP/chloroplasts/FMN pH 7.0				DCIP/chloroplasts/AQS pH 8.2			
			Maximum power (μ W)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)	Maximum power (μ W)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)
Tris-Cl	20	(+)	149	200	390	2.2	134	500	442	1.5
		(-)	117	200	380	1.9	103	500	435	0.95
	80	(+)	192	200	405	2.6	242	200	452	2.3
		(-)	151	200	395	2.0	143	600	443	1.2
Na-Tricine	20	(+)	345	200	480	3.2	159	500	452	1.6
		(-)	276	200	475	2.5	103	700	445	1.1
	80	(+)	384	100	515	3.6	232	300	458	2.2
		(-)	294	200	510	2.8	140	600	452	1.3
Phosphate	20	(+)	151	200	430	2.0	22	6,000	465	0.30
		(-)	111	300	430	1.7	2.0	900	360	0.19
	80	(+)	158	200	430	2.15	32	4,000	475	0.35
		(-)	119	300	430	1.75	2.5	900	380	0.20

FMN (3.0 mM) and AQS (1.5 mM) were used as acceptors. DCIP (0.5 mM) was the donor. (+) and (-) indicate active and inactive (killed) chloroplasts respectively.

strong reductant (reduced P430)¹³ and a weak oxidant (P700)^{5,14}. The reductant and oxidant then interact with exogenous electron acceptors and donors producing a gradient of chemical potential which can be converted into an electrical current. Also, the electron acceptor used (flavin mononucleotide, FMN) is photoreducible either directly by the electron donor (Fe^{2+}) in the case of potassium ferrocyanide¹⁵ or by Tricine¹⁶. Thus, both photosynthetic reduction of FMN mediated by PSI and FMN photochemistry (FMN photoreduction) contribute to the power developed. In the case of FMN, photodestruction also occurs. Using isolated chloroplasts in place of the PSI particles in the photochemical cell has proved advantageous because of their ease of preparation and their power conversion efficiency of close to 1%.

Isolated PSI particles or broken chloroplasts (Type C) were placed on a filter between two compartments containing the electron donor and the electron acceptor, respectively. Platinum electrodes were placed in each compartment and the photovoltage was observed across a load resistance (see ref. 12 for experimental details and design). Initially¹², we obtained an open circuit photovoltage (V_{oc}) of 360 mV, a short circuit current (I_{sc}) of 108 μ A and a maximum power (W_{max}) of 20 μ W.

We determined the characteristics of our chloroplast photoelectrochemical cell in the presence of two photoactive electron acceptors, FMN and AQS (anthraquinone 2-sulphonate) in optimal conditions of donor (dichlorophenol indophenol, DCIP) and acceptor concentrations, pH and buffer strength (Table 1). In all cases, the chloroplasts contribute at least 30% of the power observed. The two acceptors have different pH optima: 7.0 for FMN and 8.2 for AQS. Both Tris and Tricine buffers increased the power observed because they act as electron donors for the photoreduction of FMN¹⁶ and AQS¹⁷. The increase in power with increasing buffer concentration is not a

salt effect because increasing NaCl concentration from 100 to 200 mM had no effect. The AQS system did not generate greater power in phosphate buffer because, unlike FMN, there is no photodestruction which also contributes to the power developed. FMN was a better electron acceptor than AQS because of its greater absorption cross-section in the visible region of the spectrum and the fact that photodestruction could contribute to the potential developed. A maximum power of 400 μ W was observed across a 100 Ω resistance. The open circuit voltage and short circuit current were 510 mV and 3.5 mA respectively.

The time course of the photoresponse is shown in Fig. 1. A photovoltage of 485 mV was observed across a 3,000 Ω resistance. That the potential persisted for 1–2 h even after the light was turned off, illustrated the system's ability to store energy. The ability of the cell to generate a voltage is equivalent to a generator in series with a source resistance (R_s) whose magnitude is a reflection of the ability of the cell to generate current. The R_s in a photoelectrochemical cell is a function of rate of electron transport through PSI during illumination and the rate of photoreduction of the acceptor (FMN or AQS) by the reducing agents. The more efficient these processes, the lower R_s will be. Because both photosynthetic and photochemical reduction of the acceptor depend on light, R_s should be very high in dark and low in light. R_s also operates in parallel with membrane resistance (R_m). To develop and store energy, R_m has to be very high to limit the dissipation of the reduced photoproducts through the membrane. We have measured the R_m and the differences in the internal resistance in the dark and during steady state illumination (Fig. 2). The R_m was calculated to be of the order of $1.5 \times 10^5 \Omega$ (Fig. 2, curve 1). In the dark, the internal resistance was $5 \times 10^4 \Omega$ (curve 2) which decreases to 150 Ω (curve 3) after steady state photovoltage is achieved during illumination. The

Table 2 Comparison of chloroplasts and PSI particles

Donor (mM)	Acceptor (mM)	pH	Condition	Chloroplasts				PSI			
				Maximum power (μ W)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)	Maximum power (μ W)	Resistance at maximum power (Ω)	V_{oc} (mV)	I_{sc} (mA)
DCIP (0.5)	FMN (3.0)	7.0	(+)	384	100	515	4.0	432	100	521	4.4
			(-)	294	200	510	3.0	342	100	510	3.3
DCIP (0.5)	AQS (1.5)	8.2	(+)	232	300	458	2.1	306	100	460	3.3
			(-)	140	600	452	1.5	192	300	460	2.4

Chloroplasts (containing 80 μ g Chl) were aspirated onto 1-cm² filter of 1.2 μ m pore diameter. Conditions included 80 mM Na-tricine (at the indicated pH). The light intensity at the front of the cell was 5,000 W m⁻².

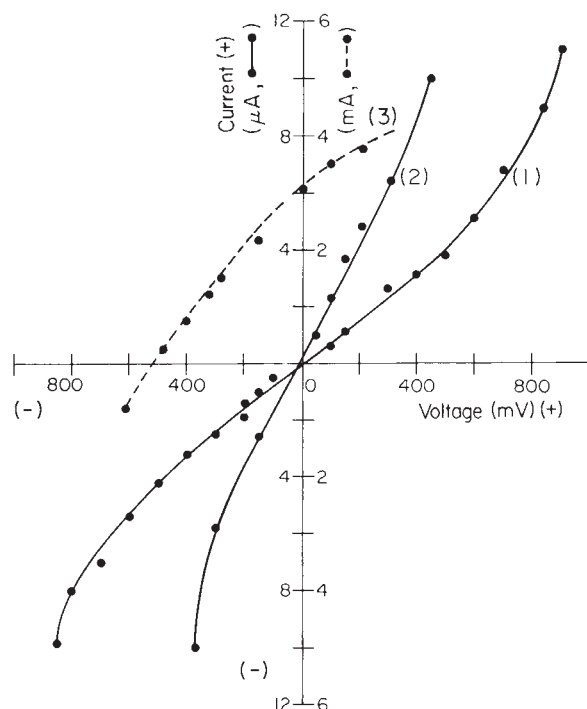


Fig. 2 Current-voltage curves of PSI impregnated membrane in absence of any dye (curve 1), in presence of dyes in dark (curve 2) and during illumination in presence of dyes after steady state potential was developed. Controlled potential was applied using a battery and a potentiometer and the current was measured. Membrane resistance (R_m) (curve 1) was measured only in salt solution. Other conditions as in Fig. 1. Light intensity was $2,000 \text{ W m}^{-2}$.

current-voltage curve shows that in light (curve 3) the V_{oc} is 520 mV and I_{sc} is 3.1 mA.

Table 2 shows that the chloroplasts can effectively substitute for the PSI particles. The power and short circuit current obtained is marginally higher in the case of PSI particles. This may be due to a permeability barrier in chloroplasts.

We calculated the power conversion efficiency for the FMN system assuming a total front surface area of 8 cm^2 . The efficiency increased as the incident light intensity was lowered from 0.08% at $5,000 \text{ W m}^{-2}$ to 0.4% for 30 W m^{-2} . The contribution of the chloroplasts to the total power also increased as the light intensity was lowered from 30% at $5,000 \text{ W m}^{-2}$ to 50% at 30 W m^{-2} . If we calculate the efficiency for 30 W m^{-2} based on the 1 cm^2 area of chloroplasts illuminated rather than the 8 cm^2 of FMN solution, we obtained an efficiency of 3.2%. The true efficiency lies between the two estimates. By subtracting the photochemical contribution from the total power an efficiency of 2.3% was obtained for the chloroplast contribution. This may not be correct, however, as the two contributions are synergistic.

This work was supported by the US Department of Energy, grant DE FG02-79ER10538. The views presented here do not necessarily reflect those of the DOE.

Received 4 August; accepted 20 November 1980.

- Bolton, J. R. in *Proc. 4th Int. Congr. on Photosynthesis* (eds. Hail, D. O. Coombs, J. & Goodwin, T. W.) 621-634 (The Biochemical Society, London, 1978).
- Tang, C. W. & Albrecht, A. C. *Nature* **254**, 507-509 (1975).
- Aizawa, M. et al. *J. Membrane Sci.* **4**, 251-259 (1978).
- Aizawa, M. et al. *Electrochim. Acta* **24**, 89-94 (1979).
- Mountz, J. M. & Tien, H. T. *Solar Energy* **21**, 291-295 (1978).
- Tien, H. T. & Mountz, J. M. *Energy Res.* **2**, 197-200 (1978).
- Tien, H. T. *Photochem. Photobiol.* **24**, 97-116 (1976).
- Allen, M. J. & Crane, A. E. *Bioelectrochem. Bioenerg.* **3**, 84-91 (1976).
- Haehnel, W. & Hochheimer, H. J. *Bioelectrochem. Bioenerg.* **6**, 563-574 (1979).
- Drachev, L. A. et al. *FEBS Lett.* **50**, 219-222 (1975).
- Janzen, A. F. & Seibert, M. *Nature* **286**, 584-585 (1980).
- Gross, E. L. et al. *Photochem. Photobiol.* **28**, 249-256 (1978).
- Ke, B. *Biochim. biophys. Acta* **301**, 1-33 (1973).
- Norris, J. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **68**, 625-628 (1971).
- Rutter, W. J. *Acta Chim. scand.* **12**, 438-446 (1958).
- Nelson, N. et al. *Photochem. Photobiol.* **16**, 481-489 (1972).
- Robinson, H. H. & Yocum, C. F. *Photochem. Photobiol.* **29**, 135-140 (1979).

Liquefaction of biomass as a source of fuels or chemicals

Ian J. Miller & S. Kevin Fellows

Chemistry Division, Department of Scientific and Industrial Research, Petone, New Zealand

Although biomass is one of the few renewable sources of liquid fuels and chemicals, relatively little work has been carried out on its direct liquefaction. Here, we report that wood or cellulose can be almost totally converted to liquids or gases when heated under pressure to 350°C in the presence of phenol, water, an acid or base catalyst, and either hydrogen or nitrogen. The resultant liquids were separated into a neutral fraction, which was partly characterized by gas chromatography-mass spectrometry (GC-MS) and found to contain mainly aromatic compounds with or without an oxygen-containing ring, and a phenolic fraction. The nature of these compounds was found to be dependent on the catalyst, but not appreciably on the physical conditions used.

Previous attempts at liquefying biomass have tended to consider it as a substitute for coal. Thus liquids have been obtained by pyrolysis under hydrogen¹, by pyrolysis in wood tars, by reaction with carbon monoxide², by hydrogenation with iodine³ or nickel⁴ as a catalyst, or by hydrogenation with a 'donor solvent'⁵. These processes, however, fail to take advantage of the depolymerization of carbohydrate fraction at least at $\sim 350^\circ\text{C}$ to monomer units such as 1,6-anhydroglucose⁶. This is followed by degradation and condensation reactions which, if unchecked, lead to the formation of char. These intermediates need therefore to react or to be rearranged before they can condense.

A preliminary objective is to slow down the condensation reactions, which is best done by providing a solvent to slow down the higher order solid state reactions. We used phenol because as the hydrogenation of lignins produces mainly phenols⁷ they would be a significant by-product of the process, and phenol with certain catalysts such as zinc chloride, dissolves cellulose. To interrupt condensation reactions we first tried hydrogenation, and then simple acid-base catalysed rearrangements.

In a typical qualitative reaction 2 g of the cellulosic material, 2 g of phenol, 2.5 g of water, and 0.5 g of other reagents were placed in a glass liner which in turn was placed in a 15 cm^3 capacity pressure vessel. This was pressurized to 3.4 MPa with either hydrogen or nitrogen, and heated to 350°C within 30 min for times between a few minutes and several hours. After cooling, the liquid products were separated into phenolic and non-phenolic fractions, and the major components of the latter characterized by GC-MS. Some of the phenolic fraction was also characterized, and was similar overall to that reported from the hydrogenation of lignin⁷.

The composition of the neutral fraction did not depend appreciably on the source of the biomass. When heated with zinc chloride and nitrogen, cellulose (filter paper, cut into squares), softwood (*Pinus radiata*, chips $\leq 6 \times 6 \times 2 \text{ mm}$) or hardwood (aspen, powdered, similarly to Masonite process, by steam explosion) gave products of almost identical composition.

The biomass was required to remain in contact with the solvent. If the volume of solvent was reduced below that outlined above, charring occurred. Only 1 g of wood chips could be used as they did not pack very well. The powdered aspen was used for most experiments because it packed better, and it could be slurried with the two phases of the liquid. To mimic green biomass, aspen with 40% dry solids, 60% moisture was used.

Liquefaction of powdered aspen was essentially complete after a few minutes at 350°C . The composition of the neutral fraction was unaffected by reaction time, although the yield was

Table 1 Liquid products obtained from aspen at 350 °C for 1 h, with phenol, water and other reagents

Reagents	% Composition of neutral products																
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
ZnCl ₂ , N ₂	8	0.6	4	1	2	62	2										
ZnCl ₂ , Ni, N ₂	0	0.6	5	1	2	58	10										
ZnCl ₂ , H ₂ , Ni	15	15	21	1		31		1*	3*	3*	0.5*	0.5*					
NaH ₂ PO ₄ , N ₂	5	4		0.5		11		7.5†					43	6			
Na ₂ CO ₃ , H ₂ , Ni	5	4		0.5		19		7.5†							5	17	5

Liquid products: (1), anisole; (2), benzofuran; (3), 2-methylbenzofuran; (4), dibenzofuran; (5), diphenylether; (6), xanthene; (7), xanthone; (8), 2,3-dimethylbenzofuran; (9), ethylbenzofuran; (10), 1,2-dimethylindane or methylindanone; (11), methylxanthene; (12), methylxanthone; (13), tetralin; (14), naphthalene; (15), cyclohexanone; (16), phenylcyclohexen-3-one; (17), ketone, tentatively phenylcyclohexan-3-one.

* Tentative identification.

† Products (8) and (9) combined.

lower at shorter times (with evidence of cellulose pyrolysates present), and at times appreciably longer than an hour, when tarring occurred. Reactions below 325 °C led mainly to char formation, and the time between 250 and 325 °C had to be minimized. This problem could be alleviated by shaking and by having larger amounts of phenol present. Reactions over 400 °C led to tarring.

Representative results using aspen are shown in Table 1. If zinc chloride was used with nitrogen, the main non-phenolic products were as listed in the first row of Table 1. If small amounts of nickel were added, there were increases in the amounts of 2-methyl benzofuran and xanthone, and a decrease of xanthene (Table 1). This may indicate that some hydrogenation occurs even without the addition of hydrogen gas, hydrogen presumably coming from the lignin. If hydrogen were used instead of nitrogen, the same products occur, but in significantly different proportions, and other compounds could be tentatively identified. About 3% of aromatic hydrocarbons, including toluene, ethyl benzene, and xylenes were also formed.

Replacing the zinc chloride with either sodium dihydrogen phosphate or sodium carbonate also gave complete liquefaction. Results and the conditions used are given in Table 1. Copper, cobalt, iron, and chromium were also used in place of nickel, and all gave liquefaction, with varying proportions of the same products.

For quantitative experiments, 2.5 g oven dried aspen, 10 g phenol, 13.5 g water, 0.8 g of a reagent and, if required 0.4 g Ni(OH)₂ were slurried, and placed in a glass liner for a larger reactor, which was pressurized to 4 MPa and heated to 350 °C in 40 min. The reagents used, and the recovered yields of neutrals were: zinc chloride/N₂, 0.76 g; zinc chloride/Ni/H₂, 0.74 g; NaH₂PO₄/N₂, 0.55 g, and K₂CO₃/N₂, 0.65 g.

Positive yields of phenol were obtained in two reactions, namely ZnCl₂/H₂/Ni (10.2 g recovered), and K₂CO₃/N₂ (10.15 g recovered). The phosphate reacted with the glass, and there was extensive char formation on the exterior of the liner. This was assumed to have come from phenolics reacting with alkaline glass, but in any case is irrelevant to further process development. That the phenol could be recycled was shown by reacting 1-inch squares of filter paper with zinc chloride, nickel, and hydrogen. The phenol was recycled three times before handling losses made it unlikely that the cellulose would dissolve properly. In these conditions 25 g of cellulose gave 12 g of neutral liquids which distilled between 50 and 400 °C.

If cellulose alone is converted to the products in Table 1 by the abstraction of oxygen atoms, the maximum yield is ~60% by weight, whereas if the conversion proceeds by the elimination of carbon dioxide, the maximum yield is ~40%. With a lignin content of 25%, the carbohydrates of aspen would give ~1.1 or 0.7 g of products if each conversion were totally efficient. The maximum yield of phenolics from the lignin would be ~0.7 g, although polyhydroxy phenols would not have been recovered in this study. Handling losses with the original solvent also prevent a proper assessment of phenol yield. The observed recoveries do establish, however, that the reaction is a net producer of phenols.

In conclusion, cellulosic material can be liquefied in conditions which, compared with those required for coal liquefaction, are relatively mild and fast. For a given set of catalysts, ~75% of the non-phenolic products consist of less than eight chemical compounds, and in some cases even fewer. Furthermore, a change in the catalysts can significantly alter the nature of the products. Thus cellulosic feedstocks could be the basis of a significant range of chemical products. Also, because hydrogen is not a necessary component of the liquefaction process for some products, it is possible that crude feedstock for a fuels or chemical industry could be manufactured in many small satellite reactors, thus obviating the large transport costs incurred if an equivalent amount of biomass were processed at one site. The satellite reactors could be low in cost, while the central processing unit could still manage the economies of scale necessary for product costs comparable with materials such as coal.

We thank Steven Gwyn for GC-MS analyses and assistance with product identification.

Received 26 August; accepted 26 November 1980.

1. Fierz-David, H. E. *Chem. Ind.* **44**, 942-4 (1925).
2. Appel, H. R., Fu, Y. C., Friedman, S., Yavorsky, P. M. & Wender, I. *U.S. Bureau of Mines* **RI 7560** (1971).
3. Geerards, J. J. Th. M., van Krevelen, D. W. & Waterman, H. I. *Brennst.-Chem.* **39**, 11-14 (1958).
4. Kaufman, J. A., Gupta, D. V., Szatkowski, T. S. & Wiess, A. H. *Chem. Eng. Tech.* **46**, 609 (1974).
5. Boomer, E. H., Argue, G. H. & Edwards, J. *Can. J. Res.* **13B**, 323-350 (1935).
6. Shafizadeh, F. *Adv. Carbohydr. Chem.* **23**, 419-474 (1968).
7. Stamm, A. J. & Harris, E. E. *Chemical Processing of Wood* Ch. 17 (Chemical Publishing Co., New York, 1953).

Endosperm cell number in barley

M. P. Cochrane & C. M. Duffus

Department of Agricultural Biochemistry, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK

Grain weight is one of the major components of grain yield and it has been suggested that in wheat, final grain weight may be a function of endosperm cell number¹. If grain weight is to be increased by increasing the rate and/or duration of endosperm cell division, it is important to establish the stage of development at which maximum endosperm cell number is attained. Evers² reported that no cell division occurred in the endosperm of wheat later than 16 days after anthesis and other workers^{3,4} have come to a similar conclusion using cell counts. However, biochemical analyses⁵ have shown that the maximum DNA content of wheat endosperm is not reached until 25 days after anthesis and both Donovan⁶ and Radley⁷ have suggested that endosperm cell division may continue until about 30 days after anthesis. We have estimated the number of endosperm nuclei in developing caryopses of barley and have examined transverse sections of endosperms of different ages. We report here that cell division in barley endosperm continues until 28-30 days after anthesis.

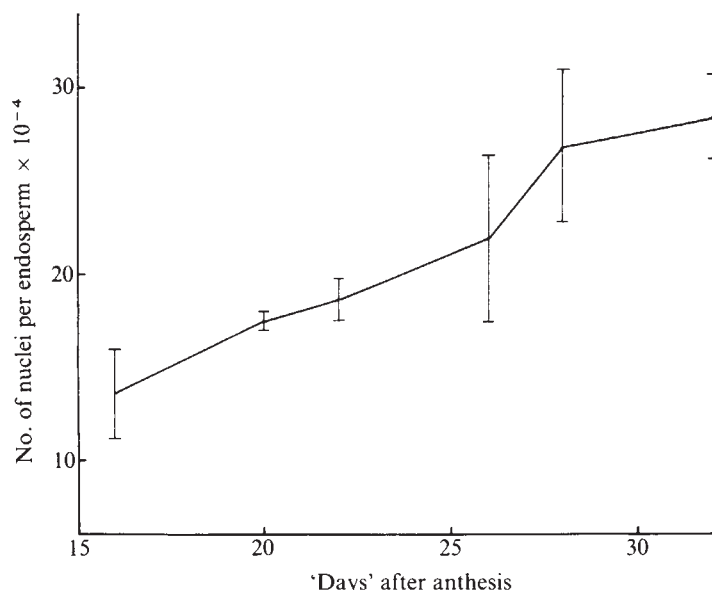


Fig. 1 The number of nuclei per endosperm during barley caryopsis development. Bars indicate confidence limits at $P = 0.05$.

In the present investigation, as in previous work in this laboratory⁸, barley (*Hordeum vulgare* var. *distichum* L. cv. Midas) was grown in a greenhouse, and caryopsis ages were estimated on a scale based on the normal development of field-grown barley in Scotland, which takes 60 days from anthesis to harvest-ripeness. The parameters used to estimate developmental age include pericarp shape and colour, endosperm shape and colour, and embryo length. The middle three or four grains were removed from each of several of the first-maturing ears in two or more pots. The developmental age of each caryopsis was determined and the endosperm dissected out, taking care to remove the testa and embryo and leave the aleurone intact. It was not possible to sample caryopses older than 33 'days' after anthesis because at this stage of development, the aleurone adheres to the testa and peels off with it. Endosperms were fixed overnight in 1:3 v/v acetic acid/ethanol and then stored in 70% ethanol. Endosperms of the same developmental age were pooled and three samples of 10 were stained using Feulgen reagent and digested in cellulase⁹. Cells were suspended in 0.5 M sorbitol and nuclei were counted in a haemocytometer. Caryopses from the same ears were prepared for light microscopy⁸.

The number of nuclei per endosperm increased during the period 16–28 'days' after anthesis (Fig. 1). Transverse sections of mid-grains (Fig. 2) show that the aleurone layer of a 15-'day' caryopsis is one to two cells thick, with cell division occurring in both anticlinal and periclinal planes. The cells in the inner and outer layers of the testa are similar in size and the cells of the nucellus are fully expanded and thin-walled. The endosperm of the 15-'day' caryopsis is flat and greyish-white, and the embryo cannot be seen with the naked eye. In 23-'day' caryopses the aleurone is two to three cells thick and cell division can be observed (Fig. 2). The cells in the outer layer of the testa are thinner than those in the inner layer, and the nucellar cells are thick-walled. At this stage the endosperm is white and rounded and the embryo is ~0.25 mm long. By 33 'days' after anthesis, three to five layers of aleurone are established and clearly differentiated from the sub-aleurone and the starchy endosperm. The inner testa cells are large and have osmiophilic contents, and the nucellus is represented only by a layer of compressed wall material. In 33-'day' caryopses the endosperm is yellowish, firm and well rounded, the embryo is ~2.1 mm long and the aleurone layer peels off with the testa.

These observations on developing barley caryopses differ from those reported for wheat in that the number of nuclei per

endosperm in barley increased up to about 30 'days' after anthesis, whereas in wheat, increase in the number of nuclei per endosperm stopped 16–20 days after anthesis^{3,4}. In addition, evidence of endosperm cell division was observed at a later stage of development in barley than that reported for wheat². These differences may be due to: (1) some assays being done on wheat endosperms contaminated with testa or stripped of their aleurone cells; (2) comparisons being made on the basis of strict chronological age, taking no account of variation in morphological development due to environmental factors; (3) the fact that the aleurone layer in wheat caryopses is only one cell thick whereas that in barley is at least three cells thick and so cell division in barley may continue for some time after the last starchy endosperm cells have been initiated. Note, however, that there is evidence of anticlinal divisions in the meristematic aleurone of 23-'day' barley endosperms and the cells thus produced presumably contribute to an increase in the surface area of the endosperm and hence to an increase in endosperm volume. It seems unlikely that similar anticlinal divisions in the aleurone of wheat cease as early as 16 days after anthesis².

Perhaps, therefore, more attention should be paid to the suggestions of Donovan⁶ and Radley⁷ that cell division in wheat endosperm continues up to 25–30 days after anthesis and to

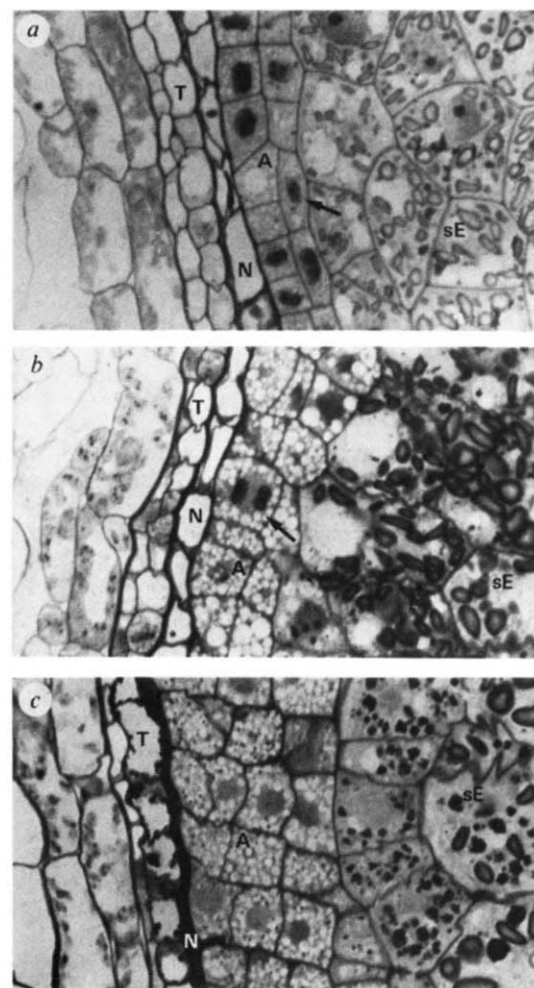


Fig. 2 Transverse sections of three barley caryopses cut at mid-grain. The material was fixed in glutaraldehyde, post-fixed in osmium tetroxide, embedded in Epon and 1.5- μ m sections were stained in toluidine blue. The tissues shown above are from the lateral part of the caryopsis. Cells undergoing division (arrows) can be seen in the aleurone layer (A) of 15-'day' and 23-'day' caryopses. These sections illustrate the state of development of the nucellus (N), testa (T) and starchy endosperm (sE) as the number of endosperm nuclei increases (Fig. 1). a, 15-'day'; b, 23-'day'; c, 33-'day'. $\times 400$.

Sandstedt¹⁰, who observed that cell divisions occurred in the endosperm of wheat, grown in Nebraska, up to 16 days after anthesis, that is, half-way between anthesis and maturity in those climatic conditions. If attempts are to be made to increase the weight of cereal grains by applying chemicals which increase endosperm cell number, it is important to establish the pattern of cell division in the developing grain. The timing of the application of any such chemicals could be particularly critical in barley because it seems that cell division taking place after a certain stage of development produces mainly aleurone cells and therefore does not provide further capacity for carbohydrate storage. Thus, stimulation of endosperm cell division early in development could increase the yield of carbohydrate, whereas stimulation of cell division at a later stage could improve the nutritional quality of the grain by increasing the yield of the proteins, fats, minerals and vitamins which are found in aleurone cells.

This work is supported by Agricultural Research Council grant AG/15/169.

Received 14 July; accepted 25 November 1980.

1. Brockelhurst, P. A. *Nature* **266**, 348–349 (1977).
2. Evers, A. D. *Ann. Bot.* **34**, 547–555 (1970).
3. Wardlaw, I. F. *Aust. J. biol. Sci.* **23**, 765–774 (1970).
4. Briarty, L. G., Hughes, C. E. & Evers, A. D. *Ann. Bot.* **44**, 641–658 (1979).
5. Donovan, G. R., Hill, R. D. & Lee, J. W. *Cereal Chem.* **54**, 546–556 (1977).
6. Donovan, G. R. *Aust. J. Pl. Physiol.* **6**, 449–457 (1979).
7. Radley, M. J. *exp. Bot.* **29**, 919–934 (1978).
8. Cochrane, M. P. & Duffus, C. M. *Ann. Bot.* **44**, 67–72 (1979).
9. Rijven, A. H. G. & Wardlaw, I. F. *Expl Cell Res.* **41**, 324–328 (1966).
10. Sandstedt, R. M. *Cereal Chem.* **23**, 337–359 (1946).

Activation of brown adipose tissue thermogenesis by the ventromedial hypothalamus

Martin N. Perkins, Nancy J. Rothwell,
Michael J. Stock & Trevor W. Stone

Department of Physiology, St George's Hospital Medical School,
Tooting, London SW17 0RE, UK

It is well established that electrical stimulation of the ventromedial hypothalamus (VMH) causes a reduction in food intake, whereas electrolytic or chemical lesions in this area result in hyperphagia and obesity in the rat. This has led to the suggestion that either the ventromedial nucleus itself, or nerve fibres passing close by, are important in the control of food intake. However, obesity due to VMH lesions occurs in weanling rats in the absence of hyperphagia and can develop in adult rats pair-fed with controls, indicating that destruction of this area also causes an increased metabolic efficiency (that is, a reduced energy expenditure). In normal rats, hyperphagia induced by feeding a highly palatable cafeteria diet is often accompanied by a large increase in heat production (diet-induced thermogenesis) which tends to prevent excessive weight gain and obesity¹. This diet-induced thermogenesis is due to sympathetic activation of brown adipose tissue (BAT) and we have now investigated the possibility that the VMH is involved in the activation of this process. We found that electrical stimulation of this area produced increased BAT thermogenesis, which suggests that the VMH exerts a dual influence in the regulation of energy balance—an inhibitory effect on energy intake and a stimulatory effect on thermogenesis and energy output.

BAT thermogenesis can be monitored by recording the tissue temperature, as Flaim *et al.*² did when stimulating the sympathetic nerves supplying interscapular BAT. This is the largest

and most accessible BAT depot and recordings can be made with minimal surgery. We therefore used the same approach to investigate the effects of electrical stimulation of the VMH on the temperature of interscapular BAT. At the same time we recorded core (rectal) and muscle (quadriceps) temperatures.

Male Sprague-Dawley rats (250–290 g) were maintained on a pelleted stock diet (PRD, Christopher Hill Group) and housed in a metabolism room at $24 \pm 1^\circ\text{C}$. They were anaesthetized with urethane (0.12 g per 100 g) and an array of three coaxial stimulating electrodes $\sim 400 \mu\text{m}$ apart in the lateral plane was stereotactically located in the VMH (De Groot's³ coordinates AP 5.4, L 0.6–0.7, V 9.0 mm), the position of all electrodes being later verified histologically. A small incision was made above the scapulae and a thermocouple probe (Comark 1604) was placed between the two lobes of the interscapular BAT and secured by closing the incision with surgical clips. A similar plastic-coated thermocouple was inserted 5 cm into the rectum and a needle thermocouple (0.5 mm o.d.) was inserted into the left quadriceps muscle. Experiments were carried out in a room kept at $24\text{--}25^\circ\text{C}$ and core temperature remained between 35.5 and 37.5°C in all animals.

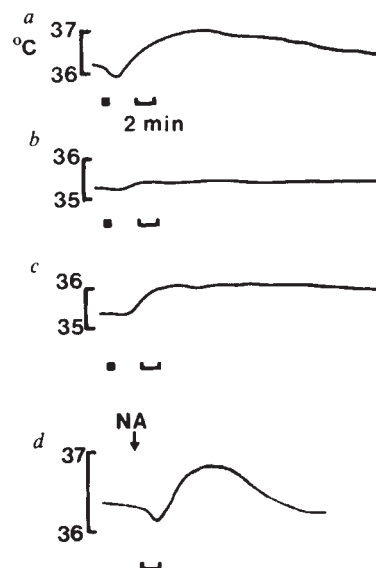


Fig. 1 Records of interscapular BAT temperature following unilateral stimulation of the VMH (■). The electrical stimulus was a 30-s train of monophasic square-wave pulses (0.5 ms, 50 Hz, 60 μA), and histological examination later confirmed that electrodes were located in the lateral and posterior regions of the VMH. *a* Shows a normal response with BAT temperature falling from 36.2 to 36.0°C 1 min after stimulation, followed by a rise to a peak value of 37.0°C 9 min later. During this period, rectal and muscle temperatures remained at 36.2 and 35.5°C , respectively. *b* Shows that treatment with the β -adrenergic antagonist propranolol (0.3 mg per kg, intravenously) reduces the rise in the BAT temperature response to less than 0.2°C . This effect of propranolol was reversible, as a normal temperature response could be elicited 2 h after treatment (*c*). *d* Shows the temperature response of BAT in a rat given an intravenous injection of noradrenaline (NA, $3 \mu\text{g}$ per kg). The initial fall in temperature (0.2°C) and the subsequent rise (0.9°C above pre-stimulation value) are similar in magnitude and duration to the response to VMH stimulation shown in *a*.

Electrical stimulation of the VMH at hourly intervals produced a biphasic response in interscapular BAT temperature (Fig. 1*a*). The initial decrease in temperature (mean $\pm$ s.e.m. = $0.16 \pm 0.05^\circ\text{C}$, $n = 5$) occurred after 1 min and was followed by a rise ($0.86 \pm 0.25^\circ\text{C}$) which was maximal at 8 ± 2 min and had returned to the pre-stimulation value after 20 ± 4 min. Rectal and skeletal muscle temperatures either did not change or increased slightly ($0.25\text{--}0.35^\circ\text{C}$) after a delay, and on each

occasion BAT temperature increased above rectal temperature by 0.3–0.5 °C. The response of BAT to VMH stimulation was remarkably similar to the effects of direct sympathetic nerve stimulation reported by Flaim *et al.*² They also noted an initial small decrease in temperature after 1 min which they ascribed to an α -adrenergic vasoconstrictor effect, and a subsequent large and prolonged increase due to β -adrenergic stimulation of BAT thermogenesis. In our study the α -antagonist phentolamine inhibited the initial decrease in temperature but did not affect the secondary increase produced by VMH stimulation. However, β -adrenergic blockade (using propranolol, Fig. 1b) almost abolished the increase in BAT temperature (0.15 ± 0.09 °C, $P < 0.01$ compared with normal response). This confirms the findings of Flaim *et al.*² and suggests that VMH activation of BAT thermogenesis also involves β -adrenergic receptors. These findings are consistent with the observation that the increase in metabolic rate of animals exhibiting non-shivering or diet-induced thermogenesis is inhibited by β -receptor blockade but less affected by α -antagonists^{1,4}.

It is difficult to denervate completely a tissue with a sympathetic supply as rich as that of BAT, and so we used a local anaesthetic to test whether activation by the VMH was mediated by neural or humoral signals. Topical application of tetracaine (1% solution) to the interscapular area abolished the response of BAT to stimulation of the VMH, although the tissue responded to injection of noradrenaline (40 μ g per 100 g intraperitoneally) with biphasic temperature responses (Fig. 1d) similar in magnitude and duration to stimulation of either sympathetic nerves² or the VMH. Central stimulation of thermogenesis by BAT therefore seems to be mediated by the sympathetic nerve supply.

The functional specificity of the VMH has been a matter of debate because, in terms of food intake, the noradrenergic fibres running through this area are apparently more important than the ventromedial nucleus⁵. This may also be true for the effects of VMH stimulation on thermogenesis but, nevertheless, the focus for successful stimulation of BAT thermogenesis can be precisely located in this area. Moving the stimulating electrodes out of the VMH by only 0.5 mm, dorsally or laterally, abolished the response of BAT to central stimulation, and normal responses could be elicited only by moving the electrodes back to their original position. Electrical stimulation in the adjacent lateral hypothalamus had no effect on BAT temperature. Others^{6,7} have noted that electrical stimulation of the VMH led to an increase in plasma glycerol and lipogenesis in BAT without affecting lipid synthesis in white adipose tissue, whereas stimulation of the lateral hypothalamus was ineffective. These findings support our suggestion that the VMH is involved in BAT thermogenesis, for diet-induced thermogenesis results in increased BAT lipolysis¹, and lipogenesis in this tissue is considerably increased during exposure to cold^{8,9}.

The importance of these findings in the regulation of energy balance depends on our proposal¹ that diet-induced thermogenesis has its origins in BAT, and we have demonstrated that the greater thermogenic capacity of hyperphagic rats fed the cafeteria diet¹ is due entirely to increased oxygen utilization by BAT¹⁰. We have also shown an insulin requirement for diet-induced thermogenesis in hyperphagic rats¹¹ which, with the results reported here, suggests that changes in glucose utilization in the insulin-sensitive VMH may be responsible for the effects of hyperphagia on diet-induced thermogenesis.

The importance of the VMH in BAT thermogenesis may explain the increased metabolic efficiency (reduced diet-induced thermogenesis) of animals with VMH lesions, particularly in young rats where obesity often develops in the absence of hyperphagia¹². This explanation is supported by the recent observation that bilateral VMH lesions result in atrophy and reduced thermogenic activity of BAT (L. Girardier, personal communication). It is likely that other areas of the hypothalamus influence BAT thermogenesis, particularly those involved in thermoregulation, but their close proximity to the VMH could provide an explanation for the strong interactions observed

between thermogenesis, food intake and environmental temperature¹³.

This work was partly supported by the Rank Prize Funds.

Received 6 August; accepted 25 November 1980.

1. Rothwell, N. J. & Stock, M. J. *Nature* **281**, 31 (1979).
2. Flaim, K. E., Horwitz, B. A. & Horowitz, J. M. *Am. J. Physiol.* **232**, R101 (1977).
3. De Groot, T. *Trans. R. Neth. Acad. Sci.* **52**, 1 (1959).
4. Schonbaum, E., Johnson, G. E., Sellers, E. A. & Gill, M. J. *Nature* **210**, 426 (1966).
5. Gold, R. M. *Science* **182**, 488 (1973).
6. Kumon, A., Takahashi, A., Hara, T. & Shimazu, T. *J. Lipid Res.* **17**, 551 (1976).
7. Shimazu, T. & Takahashi, A. *Nature* **284**, 62 (1980).
8. McCormack, J. E. & Denton, R. M. *Biochem. J.* **166**, 627 (1977).
9. Trayhurn, P. *FEBS Lett.* **104**, 13 (1979).
10. Rothwell, N. J. & Stock, M. J. *J. Physiol. Lond.* (in the press).
11. Rothwell, N. J., Stock, M. J. & Warwick, B. *Proc. Nutr. Soc.* (in the press).
12. Bernardis, L. L. & Goldman, J. K. *J. Neurosci. Res.* **2**, 91 (1976).
13. Rothwell, N. J. & Stock, M. J. *Can. J. Physiol. Pharmacol.* **58**, 842 (1980).

Photoperiodicity in the male albino laboratory rat

Edward P. Wallen

Department of Life Sciences, University of Wisconsin-Parkside, Kenosha, Wisconsin 53141

Fred W. Turek

Department of Biological Sciences, Northwestern University, Evanston, Illinois 60201

Animals inhabiting areas where there are drastic changes in the environment often reproduce only during limited time periods to ensure that young are raised in optimal environmental conditions¹. The lack of a well defined breeding season in many domesticated animals^{2,3}, presumably because the selective pressures for seasonal breeding have been minimized, suggests that the neuroendocrine events controlling seasonal cyclicity have been bred out of these animals. Little is known about the underlying neuroendocrine changes which may occur during the evolution of a species from a seasonal to a nonseasonal breeder. Whereas the changing photoperiod is the primary environmental cue which initiates and/or terminates the reproductive season in many animals⁴, this is not so in the albino rat *Rattus norvegicus*, a model nonseasonal breeder^{5–7}. Nevertheless, daylength can influence various reproductive parameters in laboratory rats^{8–17}, suggesting that some of the neuroendocrine components that controlled seasonal breeding previously are still extant in this species. To test this hypothesis, we investigated the effect of daylength on the responsiveness of the neuroendocrine–gonadal axis to the negative-feedback effects of testosterone. This paradigm was chosen because of the important role played by photic-induced changes in steroid feedback sensitivity in the control of seasonal reproduction^{18–21}. We report here that although daylength has very little effect on neuroendocrine–gonadal function in the intact male laboratory rat, it seems that some component(s) of a photoperiodic system involving the pineal gland has been preserved.

Male inbred Sprague–Dawley albino rats (Gibco, 180–200 g) were housed (four per cage) with food (Purina Rat Chow) and water provided *ad libitum*. Lighting was provided by banks of General Electric cool-white 40-W fluorescent bulbs. The animals were maintained on a light/dark 14:10 h cycle for a 2-week period before starting the experiment. They were then either pinealectomized²² or sham pinealectomized and implanted subcutaneously with either empty or testosterone-filled Silastic capsules that were 10 or 15 mm long (16 animals per group). Half of the animals in each group remained on LF14:10 while the remaining animals were transferred to LD 6:18. Sixty-three days later the animals were killed by decapitation, blood was collected and the testes removed and weighed. Serum was assayed for luteinizing hormone (LH) and follicle-stimulating hormone (FSH) by methods described previously²³.

Consistent with the idea that the male albino laboratory rat is not a photoperiodic species is the observation that testicular weight and serum levels of LH and FSH were similar in untreated animals exposed to either LD 6:18 or 14:10 (Fig. 1, Table 1). On the other hand, in intact rats implanted with 10- or 15-mm long testosterone-filled capsules for 63 days, the testes were significantly ($P < 0.05$) smaller in animals exposed to LD 6:18 compared with those of animals maintained on LD 14:10. The degree of testicular regression observed in the LD 6:18 animals implanted with a 15-mm long testosterone-filled capsule was similar to the maximum regression observed in a previous dose-response study using various sizes of Silastic capsules filled with testosterone²⁴.

Pinealectomy had no effect on testicular weight in untreated animals exposed to either photoperiod, or in testosterone-treated animals exposed to LD 14:10 (Fig. 1). In contrast, pinealectomy abolished the increased responsiveness to the inhibitory effects of testosterone in animals exposed to short days. Pinealectomized animals maintained on LD 6:18 and implanted with either 10- or 15-mm long testosterone-filled capsules had significantly larger testes ($P < 0.05$) than intact rats on LD 6:18 that were implanted with the same size capsules.

The decrease in testicular weight observed in testosterone-treated rats (Fig. 1) was presumably due to an inhibition of pituitary gonadotropin release²⁵. Serum FSH levels were significantly reduced ($P < 0.05$) in all of the testosterone-treated pineal-intact animals when compared with the levels found in animals receiving empty implants. Although serum levels of LH were consistently lower in the pineal-intact rats treated with testosterone than in the untreated control animals, these differences were not statistically significant. However, note that serum LH levels in the intact untreated rats were already near the lower limits of detectability of our assay, and a significant decrease would be difficult to observe.

Surprisingly, no significant effects of either the photoperiod or pinealectomy on serum levels of FSH and LH were observed in testosterone-treated rats (Table 1). This occurred despite the fact that both the photoperiod and pinealectomy did alter the effect of testosterone on reducing testicular size. It is not known how an equal dose of testosterone (that is, 10- or 15-mm testosterone-filled capsules) was able to reduce testis size differentially without also having a differential effect on serum FSH and LH levels. Possibly, a more complete sampling regimen would have yielded differences in gonadotropin levels which were not demonstrable using a single terminal sample. Alternatively, the effect of testosterone may have been via some other hormone which influenced testicular activity in response to the photoperiod and/or pinealectomy. Prolactin is a likely candidate as previous studies indicate that prolactin may be important in the photoperiodic control of testicular activity in rodents^{26,27}.

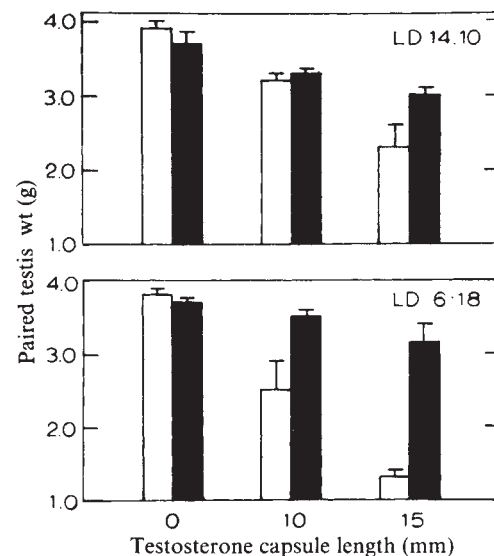


Fig. 1 Mean paired testis weight of rats that were either sham pinealectomized (open bars) or pinealectomized (shaded bars) and implanted with either empty or testosterone-filled Silastic capsules that were 10 or 15 mm long. At the time of capsule implantation the animals were either maintained on LD 14:10 (top panel) or were transferred to LD 6:18 (bottom panel) for 63 days. The vertical line represents the standard error. The testosterone-filled capsules were prepared using Silastic tubing (Dow Corning). Approximately 40 µg of testosterone are released per day per 10 mm of capsule length.

Recent studies on two photoperiodic species, the golden hamster and the sheep, demonstrate the importance of a photic-induced alteration in the responsiveness of the hypothalamic-pituitary axis to the negative-feedback effects of sex steroids in the control of seasonal breeding^{18-21,23}—exposure to a nonstimulatory photoperiod renders the animals extremely sensitive to the negative-feedback effects of steroid hormones on pituitary gonadotropin release. Furthermore, in the golden hamster pinealectomy can block the hypersensitivity to testosterone feedback induced by short days²⁸. That pinealectomy can block the short day-induced increase in responsiveness to testosterone in both the albino rat and the hamster suggests that the neuroendocrine events associated with photo-induced changes in sensitivity are similar. Indeed, the photoperiod can alter the negative feedback effect of oestrogen on pituitary gonadotropin release in the ovariectomized albino rat⁸⁻¹⁰. However, in the present study, serum levels of LH and FSH were similar in the testosterone-treated pineal-intact rats exposed to LD 14:10 and LD 6:18. Further studies are required to determine how photic-induced changes in sensitivity to steroid hormones in seasonal breeders are related to the change observed in the non-seasonally breeding laboratory rat.

There are conflicting reports regarding the effect of both the photoperiod and pinealectomy on reproductive function in the rat^{5-17,29-34}. The present study suggests that such apparent discrepancies may be due to some components of the neuroendocrine system which may have been involved in the photoperiodic control of reproduction at a previous evolutionary time still existing in the rat, and being manifested only in certain experimental conditions. An added complication is that the inbreeding of various strains of laboratory rats may have selected for or against different components of the photoperiodic-neuroendocrine system.

The results reported here on the albino rat are paradoxical. If short days render the animals more sensitive to the inhibitory effects of testosterone, one would expect that testicular regression would occur in untreated rats exposed to short days, as these animals would be more responsive to the testosterone produced by their own testes. One explanation for our results is that the inhibitory effects of short days are blocked or offset by other short day-induced changes which occur in the

Table 1 Measured (mean $\pm$ s.e.m.) serum levels of LH and FSH (ng ml^{-1}) in intact and pinealectomized rats

Testosterone capsule size (mm)		LD 14:10		LD 6:18	
		Intact	Px	Intact	Px
0	LH	1.2 $\pm$ 0.4	1.2 $\pm$ 0.3	1.5 $\pm$ 0.4	1.9 $\pm$ 0.2
	FSH	336 $\pm$ 27	302 $\pm$ 17	301 $\pm$ 18	345 $\pm$ 22
10	LH	0.5 $\pm$ 0.2	1.0 $\pm$ 0.2	0.6 $\pm$ 0.2	0.6 $\pm$ 0.1
	FSH	226 $\pm$ 20	277 $\pm$ 10	220 $\pm$ 25	308 $\pm$ 37
15	LH	0.6 $\pm$ 0.2	0.4 $\pm$ 0.1	0.3 $\pm$ 0.0	0.9 $\pm$ 0.4
	FSH	236 $\pm$ 15	289 $\pm$ 52	194 $\pm$ 15	246 $\pm$ 2

The rats were implanted with either empty or testosterone-filled capsules of 10- or 15-mm length. After capsule implantation, the animals were exposed to either LD 14:10 or LD 6:18 for 63 days. LH and FSH values are expressed in terms of nanogram equivalents of NIH-LH-S16 or NIAMMD-Rat-FSH-RP-1.

Px, pinealectomized.

neuroendocrine system. This hypothesis is particularly attractive because recent evidence demonstrates that at least two mechanisms may be involved in the photoperiodic response of seasonal breeders. In addition to the effects of daylength on steroid-feedback sensitivity, it seems that the photoperiod can also influence pituitary gonadotropin release in the absence of gonadal steroid hormones³⁵⁻³⁸. In rats exposed to short days, there may be neuroendocrine changes which are independent of steroid feedback and which negate the enhanced negative-feedback effects of steroid hormones. Because the photoperiod

alters the neuroendocrine-gonadal axis only in certain conditions, the laboratory rat may be useful for the elucidation of the neuroendocrine components involved in the photoperiodic control of reproduction in seasonal breeders.

We thank Sue Losee, John Georgeson and John Smith for technical assistance, Susan Stice for preparing the manuscript, and Gary Ellis for his comments and criticism. This work was supported by NSF grant PCM-7823377 to E.P.W. and NIH grants HD-09885, HD-12622 and Research Career Development Award HD-00249 to F.W.T.

Received 4 August; accepted 21 November 1980.

- Assenmacher, I. & Farner, D. S. (eds) *Environmental Endocrinology* 1-122 (Springer, Berlin, 1978).
- Lodge, J. R. & Salisbury, G. W. in *The Testis* Vol. 3 (eds Johnson, A. D., Gomes, W. R. & Vandemark, N. L.) 139-167 (Academic, New York, 1970).
- Schwartz, N. B. in *The Hypothalamus* (eds Martini, L., Motta, M. & Fraschini, F.) 515-528 (Academic, New York, 1970).
- Turek, F. W. & Campbell, C. S. *Biol. Reprod.* **20**, 32-50 (1979).
- Kinson, G. A. & Robinson, S. J. *J. Endocr.* **47**, 391-392 (1970).
- Reiter, R. J. *et al. Endocrinology* **88**, 895-900 (1971).
- Kinson, G. A. & Liu, C. *Life Sci.* **14**, 2179-2188 (1974).
- Hoffmann, J. C. *Biol. Reprod.* **2**, 255-261 (1970).
- Hoffmann, J. C. & Cullin, A.-M. *Neuroendocrinology* **23**, 285-296 (1977).
- Wallen, E. P. & Yochim, J. M. *Biol. Reprod.* **11**, 117-124 (1974).
- McCormack, C. E. *Fedn Proc.* **31**, 811 (1972).
- Sridaran, R. & McCormack, C. E. *Biol. Reprod.* **20**, 705-712 (1979).
- Yochim, J. M. & Mitchell, J. A. *Endocrinology* **87**, 465-471 (1970).
- Mitchell, J. A. & Yochim, J. M. *Endocrinology* **87**, 472-480 (1970).
- Sorrentino, S. J. & Benson, B. *Gen. comp. Endocr.* **15**, 242-246 (1970).
- Shapiro, M. I., Collu, R., Masse, D., Ducharme, J. R. & Roux, J. F. *Life Sci.* **19**, 1341-1346 (1976).
- Piacsek, B. E. & Meites, J. *Neuroendocrinology* **2**, 129-137 (1967).
- Pelletier, J. & Ortavant, R. *Acta endocr.* **78**, 442-450 (1975).
- Tamarkin, L., Hutchison, J. S. & Goldman, B. D. *Endocrinology* **99**, 1528-1533 (1976).
- Legan, S. J., Karsch, F. J. & Foster, D. L. *Endocrinology* **101**, 818-824 (1977).
- Ellis, G. B. & Turek, F. W. *Endocrinology* **104**, 625-630 (1979).
- Hoffman, R. A. & Reiter, R. J. *Anat. Rec.* **153**, 19-22 (1965).
- Turek, F. W. *Endocrinology* **101**, 1210-1215 (1977).
- Berndtson, W. E., Desjardins, C. & Ewing, L. L. *J. Endocr.* **62**, 125-135 (1974).
- Campbell, C. S. & Schwartz, N. B. *J. Tox. envir. Hlth* **3**, 61-95 (1977).
- Bex, F. J. & Bartke, A. *Endocrinology* **100**, 1223-1226 (1977).
- Bartke, A. *et al. Endocrinology* **106**, 167-172 (1980).
- Turek, F. W. *Endocrinology* **104**, 636-640 (1979).
- Ronnekleiv, O. K. & McCann, S. M. *Neuroendocrinology* **19**, 97-114 (1975).
- Alonso, R., Prieto, L. & Mas, M. *Endocrinology* **102**, 1534-1538 (1978).
- Jacobson, C. D. & Mann, D. R. *Biol. Reprod.* **18**, 712-718 (1978).
- Brown-Grant, K. & Ostberg, A. J. C. *J. Endocr.* **62**, 45-50 (1974).
- Blake, C. A. *J. Endocr.* **69**, 67-75 (1976).
- Takeo, Y., Shirama, K., Shimizu, K. & Maekawa, K. *Annotes zool. jap.* **49**, 83-89 (1976).
- Davis, G. J. & Meyer, R. K. *Gen. comp. Endocr.* **20**, 61-68 (1973).
- Gibson, W. R., Follett, B. K. & Gledhill, B. J. *Endocr.* **64**, 87-101 (1975).
- Pelletier, J. & Ortavant, R. *Acta endocr.* **78**, 435-441 (1975).
- Ellis, G. B. *Biol. Reprod.* **20**, Suppl. 1, 30A (1979).

A human quantitative polymorphism related to Xg blood groups

P. N. Goodfellow

Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

P. Tippett

MRC Blood Group Unit, Wolfson House, University College, London NW1 2HE, UK

A monoclonal antibody, 12E7, raised against lymphocytes from a patient with a T-cell acute lymphocytic leukaemia reacts strongly with cortical thymocytes and, to a lesser extent, with other human cells¹. The antigen defined by 12E7 is not expressed on mouse or hamster cells; this species specificity allowed us to investigate the genetics of the expression of the 12E7 antigen using human-rodent somatic-cell hybrids². Hybrid cells which contain the human X chromosome as the only human genetic contribution react with 12E7. The X-linkage of the gene, or genes, controlling 12E7 antigen expression and the fact that 12E7 antigen is also expressed on red blood cells led us to investigate the possibility of a relationship with the X-linked Xg blood group system. Although the 12E7 and Xg^a antigens seem to be different antigenic determinants specified by different loci, we report here a sex-limited effect involving the Xg phenotype on the expression of 12E7 antigen.

Expression of the Xg blood group system is controlled by two X-borne alleles, Xg^a and a silent allele Xg. The antigenic product, Xg^a, of the dominant Xg^a allele is recognized by alloantisera using the indirect antiglobulin haemagglutination method^{3,4}. Experiments to demonstrate Xg^a antigen on tissues other than red blood cells have given conflicting results^{5,6}. A quantitative polymorphism of the 12E7 antigen on red blood cells was first detected in a family segregating at the Xg locus.

The mother J. B., the father W. B. and daughter H. B. all express the Xg^a antigen—this phenotype is designated Xg(a+). The son M. B. has the alternative Xg(a-) phenotype which lacks the Xg^a antigen, indicating that the mother J. B. is heterozygous Xg^aXg. The Xg(a-) son showed a reduced expression of the 12E7 antigen compared with other family members (Fig. 1). The same quantitative polymorphism of 12E7 antigen expression

Table 1 Results of tests for 12E7 antigen phenotype in 10 random donors

Donor	c.p.m. × 10 ⁻³			Antiglobulin test score	12E7 phenotype
	12E7	W6/34	CA2.06	12E7	
1	24.0	31.9	2.0	30	High
2	9.1	32.6	1.1	13	Low
3	28.5	33.5	1.2	30	High
4	23.6	37.2	1.1	30	High
5	37.8	37.1	2.2	28	High
6	35.8	42.5	1.4	28	High
7	9.2	33.8	1.1	13	Low
8	34.1	37.4	1.1	30	High
9	36.6	34.4	1.8	30	High
10	8.6	37.6	1.3	0	Low

The binding assay was carried out as described in Fig. 1 legend. 12E7 was used at a dilution of 1:200. W6/34, a monoclonal antibody which reacts with a human red-cell antigen, at 1:200 dilution, was the positive control. CA2.06, the negative control, was used as described in Fig. 1 legend. The low phenotype was scored if the counts bound with 12E7 were less than half those bound with W6/34 in the same conditions. Repeated testing of several individuals gave consistent results. For phenotypic scoring of 12E7 by the antiglobulin haemagglutination method, see Table 2 legend.

was also found when red blood cells from random donors were tested (Table 1).

Although the 12E7 antigen was originally detected by the indirect radioimmunoassay (RIA), it could also be demonstrated by an indirect antiglobulin haemagglutination test with the same pattern of high and reduced expression (Table 1); over 50 people were typed by both techniques and complete concordance between the different methods was achieved. Subsequent tests were carried out using the haemagglutination method.

The relationship between the Xg locus and the 12E7 polymorphism was investigated by testing over 300 Europeans for Xg^a and for the level of expression of 12E7 antigen (Table 2). All Xg(a+) individuals show high levels of 12E7 antigen expression but Xg(a-) individuals fall into two categories, showing either high- or low-level 12E7 phenotype. The probability that these results occurred by chance is less than 1 in 10⁸, indicating a relationship between the Xg locus and the 12E7 polymorphism. The simplest explanation for these results is that the 12E7

polymorphism is controlled by a previously undetected variant of the silent Xg allele. Thus, the Xg^a allele and one Xg allele would lead to a high-level expression of 12E7; a different Xg allele would give low-level expression of the 12E7 antigen. Alternatively another polymorphic locus might exist with an allele for low-level expression of the 12E7 antigen in very strong linkage disequilibrium with the Xg allele. However, both these explanations fail to explain the sex distribution shown in Table 2b. Although $Xg(a-)$ males may be high- or low-level expressors of 12E7, $Xg(a-)$ females always have the low-level 12E7 phenotype ($P < 0.001$).

These data imply sex limitation of 12E7 expression. One hypothesis is that there is a polymorphic 12E7 modifying gene and that the allele for high-level 12E7 expression cannot be expressed in females lacking the Xg^a antigen.

An alternative, and perhaps controversial, hypothesis is that a locus, Yg , equivalent to Xg , is present on the Y chromosome. Like Xg , the Yg locus must have two alleles, Yg^a and Yg . It is postulated that Xg^a and Yg^a allow high-level expression of the 12E7 antigen, whereas Xg and Yg do not. Thus all $Xg(a+)$ individuals ($Xg^a Xg^a$, $Xg^a Xg$ females and $Xg^a Yg^a$, $Xg^a Yg$ males) and $Xg(a-)$ $Yg(a+)$ males ($Xg Yg^a$) express high levels of 12E7, whereas $Xg(a-)$ females ($Xg Xg$) and $Xg(a-)$ $Yg(a-)$ males ($Xg Yg$) express a low level of 12E7. It is interesting that the gene frequency of the postulated Yg^a allele (0.68), as calculated from $Xg(a-)$ males who are 12E7 high expressors (Table 2b), is very similar to that of the Xg^a allele (0.66) in similar populations⁵. We are now searching for the rare families

Table 2 Two by two tables comparing Xg^a and 12E7 expression

a		$Xg(a+)$	$Xg(a-)$
	12E7 expression		
	High	237	56
	Low	0	42
b		$Xg(a-)$	
	12E7 expression	♂	♀
	High	56	0
	Low	26	16

Phenotypic scoring of 12E7 was carried out 'blind' using either the indirect antiglobulin technique or both indirect RIA and the antiglobulin method. For the latter, washed red blood cells sensitized with 12E7 antibody were tested with rabbit anti-mouse globulin by the capillary method and agglutination was scored at 10, 20 and 30 min. High expression: 25–30; low expression: 0–18. Male $Xg(a-)$ with low expression generally scored more, 0–18, than did female $Xg(a-)$, 0–5.

which can be used to distinguish between these hypotheses and are attempting to investigate the biochemical relationship between the Xg^a and 12E7 antigens.

We thank our colleagues who gave blood for this study; Dr M. Contreras and Mr P. Bowell for providing many blood samples and Drs P. W. Andrews, J. Bodmer, W. F. Bodmer, B. Knowles, R. Sanger, E. Solomon and J. Trowsdale for discussion and guidance. The 12E7 monoclonal antibody was produced and generously donated by Dr R. Levy.

Received 3 September; accepted 18 November 1980.

1. Levy, R., Dilley, J., Fox, R. I. & Warnke, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6552–6556 (1979).
2. Goodfellow, P., Banting, G., Levy, R., Povey, S. & McMichael, A. *Somat. Cell Genet.* **6**, 777–787 (1980).
3. Mann, J. D. *et al. Lancet* **i**, 8–10 (1962).
4. Race, R. R. & Sanger, R. *Blood Groups in Man* 6th edn (Blackwell Scientific, Oxford, 1975).
5. Fellous, M., Bengtsson, B., Finnegan, D. & Bodmer, W. F. *Ann. hum. Genet.* **37**, 421–430 (1974).
6. Hsu, H. S., Migeon, B. R. & Bias, W. B. *Human Gene Mapping 3, Birth Defects Ser.* **12**, Vol. 7, 382–385 (1976).
7. Charron, D. J. & McDevitt, H. O. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6567–6571 (1979).

Genetic basis of BCG-induced suppression of delayed hypersensitivity

Denis J. Schrier, John L. Sternick, Elizabeth M. Allen & Vernon L. Moore

Research Service, Wood Veterans' Administration Medical Center, and the Departments of Medicine and Pathology, The Medical College of Wisconsin, Milwaukee, Wisconsin 53193

BCG can either act as an adjuvant to potentiate immunological responses or, in some cases, can induce suppression. The reasons for these differential activities are not clear but may include routes and doses of administration, as well as variable host reactivity to the agent. In this study, we have used killed BCG administered intravenously to produce chronic granulomatous inflammation (CGI) in the lungs and spleen of inbred mice. We report that strains which developed CGI were usually anergic, as evaluated by the development of delayed hypersensitivity (DH) to sheep erythrocytes (SRBC). Studies on the genetics of BCG-induced anergy indicated that it was unigenic, recessive and linked (approximately 28 recombination units) to the immunoglobulin heavy-chain allotype (*Igh*). There was no influence by genes linked to the major histocompatibility complex. The study indicates that anergy associated with CGI is under genetic control, which may explain the variability of anergy in patients with granulomatous diseases. The implication of linkage to the *Igh* complex is not clear, but it may be associated with V_H receptors on T lymphocytes, which in turn act on macrophages to mediate suppression.

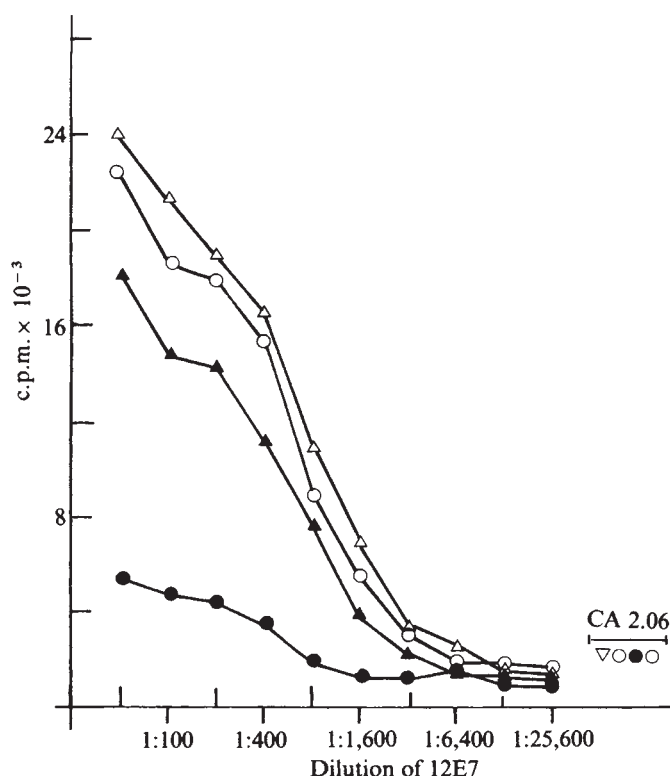


Fig. 1 Binding of 12E7 to red blood cells from a family segregating at the Xg locus. 12E7 tests on red cells of: $\triangle$, father (W.B.); $\circ$, daughter (H.B.); $\blacktriangle$, mother (J.B.); $\bullet$, son (M.B.). The indirect RIA was as described elsewhere². Red blood cells carried out (5×10^6) were mixed with 50 μ l of antibody and incubated overnight at 4 °C. The cells were washed and 2×10^5 c.p.m. of 125 I-labelled, column-purified, rabbit anti-mouse immunoglobulin added. The mixture was incubated at 4 °C for 1 h. The cells were washed and the bound second antibody quantified in a gamma counter. All dilutions and washes were performed with Dulbecco's phosphate-buffered saline supplemented with 1% bovine serum albumin and 0.1% sodium azide. The 12E7 antibody was prepared by ammonium sulphate precipitation from the cultured hybridoma supernatant². The negative control, CA2.06, a monoclonal anti-human DRw, was used as undiluted culture supernatant⁷.

We are interested in BCG because it can be used to induce CGI in the lungs of rabbits¹ and in the lungs and spleens of only certain strains of inbred mice². A seemingly paradoxical observation has been that inbred mice which develop CGI in response to BCG are suppressed in their response to mitogens and in the development of DH and a primary antibody response to SRBC³. Mice which did not develop CGI were not suppressed in any of these reactivities. Suppression was mediated by cells from inflamed spleens which were adherent to plastic petri plates, lacked Thy 1 and immunoglobulin surface markers, and were removed with carbonyl iron⁴. These traits are characteristic of macrophages.

We examine here the genetic basis of BCG-induced suppression in various inbred mice, in their F₁, F₂ and backcross progeny, and in selected congenic strains. Data in Table 1 indicate that whereas C57BL/6 (B6) mice (which develop CGI in response to BCG) are markedly suppressed in their ability to develop DH to SRBC, CBA mice (which do not develop CGI to BCG)² express DH to SRBC. This is not due to an innate inability of B6 mice to develop DH to SRBC as both strains respond when not treated with BCG (Table 2). Table 1 also indicates that BCG-induced suppression is recessive because BCG-treated (B6×CBA)F₁ mice were not suppressed. Analysis of F₂ progeny of (B6×CBA)F₁ mice showed that 26.6% (33/124) were suppressed. Assuming independent segregation, these data suggest that a single gene or gene complex is responsible for suppression. Note also that several of the F₂ progeny of (B6×CBA)F₁ mice which developed BCG-induced anergy displayed only minimal inflammation. On the other hand, many of the F₂ progeny which were not anergic developed intense inflammation. Although this suggests that genes controlling the development of CGI are different from those which influence anergy, we realise that this is difficult to prove in this system because of the multigenic control of CGI, as contrasted with single gene control of anergy⁵. Because F₁ offspring of either sex from B6 females are not suppressed (data not shown), the control of suppression does not seem to be sex-linked.

To test whether suppression was linked to the *H-2* complex, we performed linkage analysis in F₂ progeny of (B6×CBA)F₁ parents. The F₂ progeny were typed for either the *H-2^b* or *H-2^k* allele using anti-*H-2* serum prepared in our laboratory. Anti-*H-2^k* serum was produced by injecting B10.BR cells into C57BL/10 mice and anti-*H-2^b* serum was produced by injecting C3H.SW cells into C3H mice. The following results were

Table 1 BCG-induced suppression of delayed hypersensitivity to sheep erythrocytes

Mouse strain	Footpad increment (mm)*		No. suppressed total (%)
	Positive	Negative	
C57BL/6		0.16±0.02	8/8 (100)
CBA	0.77±0.06		0/10 (0)
[B6×CBA]F ₁	0.79±0.04		0/22 (0)
[(B6/CBA)×(B6/CBA)]F ₂	0.61±0.02	0.17±0.01	33/124 (26.6)†

Mice were injected intravenously with 300 µg of killed BCG in an oil-in-saline emulsion. Four weeks later, B6, but not CBA, mice exhibited intense CGI in their lungs and spleen.

* Increase in 24-h footpad swelling due to specific antigen compared to footpad injected with equal volume of saline. Footpad values >0.31 mm are considered positive; this figure was obtained by calculating two standard deviations from the mean footpad increment of animals developing positive DH responses. DH was produced by injecting mice intravenously with 5×10⁵ SRBC 4 weeks after BCG injection. Five days later, animals were challenged in one rear footpad with 10⁸ SRBC and in the other hind footpad with an equal volume of saline. Increases in footpad thickness were quantified using a pressure-sensitive caliper (Dyer).

† 33/124 (26.6%) is not different from an expected percentage of 25 (*P*>0.5) predicted in the F₂ generation for a recessive trait controlled by a single gene.

Table 2 Strain survey of BCG-induced suppression of delayed hypersensitivity to sheep red blood cells

Mouse strain	BCG injected	<i>H-2</i>	<i>Igh</i>	Footpad increment*
C57BL/6	+	<i>b</i>	<i>b</i>	0.16±0.02†
	—			0.60±0.06
CBA	+	<i>k</i>	<i>a</i>	0.77±0.06
	—			0.57±0.06
C57BL/10	+	<i>b</i>	<i>b</i>	0.17±0.06†
	—			0.68±0.09
C57BR	+	<i>k</i>	<i>a</i>	1.12±0.16
	—			0.41±0.05
C57L	+	<i>b</i>	<i>a</i>	1.16±0.09
	—			1.24±0.26
BALB/c	+	<i>d</i>	<i>a</i>	0.51±0.08
	—			0.46±0.05
BALB/c. <i>Igh^b</i>	+	<i>d</i>	<i>b</i>	0.18±0.02†
	—			0.60±0.06
C3H.SW	+	<i>b</i>	<i>a</i>	0.48±0.06
	—			0.41±0.04
B10.BR	+	<i>k</i>	<i>b</i>	0.14±0.03†
	—			0.59±0.08

* See Table 1 legend.

† Negative footpad responses.

obtained: *H-2^{b/b}*, suppressed (10); *H-2^{b/b}* nonsuppressed (18); *H-2^{b/k}* suppressed (11); *H-2^{b/k}* nonsuppressed (40); *H-2^{k/k}* suppressed (4); and *H-2^{k/k}* nonsuppressed (18). These results do not differ significantly from a 6.25:18.75:12.5:37.5:6.25:18.75 ratio based on independent segregation of a recessive trait (*P*>0.10). A strain distribution analysis (Table 2) also suggested that genes linked to the *H-2* complex are not involved in suppression. In particular, C57BL/6 (*H-2^b*), C57BL/10 (*H-2^b*), B10.BR (*H-2^k*), and BALB/c.*Igh^b* (*H-2^d*) all exhibited suppression even though they were different *H-2* types. Table 2 shows also that all strains which were suppressed were *Igh^b*, suggesting linkage to the *Igh* allotype. This is further substantiated because BALB/c mice (*Igh^a*) were not suppressed, but the congenic partner, BALB/c.*Igh^b*, was. We analysed linkage to the *Igh* complex further by performing linkage analysis on male and female (C57BL/6×CBA)F₁×C57BL/6 backcross mice. The backcrosses were typed for either the *Igh^b* or *Igh^a* allele with the following results: *Igh^{b/a}*, nonsuppressed (14); *Igh^{b/b}*, suppressed (9); *Igh^{b/a}*, suppressed (6); *Igh^{b/b}*, nonsuppressed (3). These results differ significantly from a 1:1:1:1 ratio expected on the basis of independent segregation (*P*<0.025). The estimated recombination frequency is 0.28±0.08.

In summary, BCG-induced suppression of DH to SRBC is under genetic control. It is recessive, autosomal, due to a single (or closely linked group of) gene, and controlled by genes linked to the *Igh* complex.

Macrophage-mediated suppression is caused by a variety of agents, is usually not immunospecific, and suppresses a variety of immune responses⁶⁻⁸. Its overall role in immunoregulation is not understood. In our system, this type of suppression occurs in conjunction with intense CGI in the lungs and spleen. In another study, we have shown that adherent spleen cells from BCG-treated B6 mice produce soluble factors *in vitro* that suppress the production of CGI in the lungs and spleen⁹. Therefore, macrophage-mediated suppression may be important in the control of inflammation. Further studies are needed, however, to test this hypothesis.

The significance of linkage to *Igh* allotypes is not yet clear. In some studies, macrophage-mediated suppression is controlled by T lymphocytes^{10,11}, and several T cell-related immunological phenomena are linked to the *Igh* allotype^{12,13}, which may be related to V_H receptors on T cells. Studies are in progress in our laboratory to test these possibilities.

This research was supported by funds from HL22301, HL15389 and the Veterans Administration. We acknowledge

the BALB/c.*Igh^b* mice from Dr Noel Warner, a gift of anti-*Igh* serum by Dr Mel Bosma, and the editorial assistance of Ms Mavis Baurle and Peggy LaSusa.

Received 19 February; accepted 27 October 1980.

- Galindo, B. & Myrvik, Q. N. *J. Immun.* **105**, 227–237 (1970).
- Allen, E. M., Moore, V. L. & Stevens, J. O. *J. Immun.* **119**, 343–347 (1977).
- Allen, E. M. & Moore, V. L. *J. reticuloendothel. Soc.* **26**, 349–356 (1979).
- Schrier, D. J., Allen, E. M. & Moore, V. L. *Fedn Proc.* **38**, 1174 (1979).
- Moore, V. L., Schrier, D. J., Sternick, J. L. & Allen, E. M. in *Basic and Clinical Aspects of Granulomatous Inflammation* (eds Boros, D. L. & Yoshida, T.) (Elsevier, New York, in the press).
- Coulis, P. A., Lewert, R. M. & Fitch, F. W. *J. Immun.* **120**, 1074–1076 (1978).
- Klimpel, G. R. & Henney, C. S. *J. Immun.* **120**, 563–569 (1978).
- Kirchner, H., Chused, T. M., Herberman, R. B., Holden, H. T. & Lavrin, D. H. *J. exp. Med.* **139**, 1473–1487 (1974).
- Allen, E. M., Calvanico, N. J. & Moore, V. L. *Fedn Proc.* **38**, 1174 (1979).
- Tadakuma, T. & Pierce, C. W. *J. Immun.* **120**, 481–486 (1978).
- Ptak, W., Zembala, M. & Gershon, R. K. *J. exp. Med.* **148**, 424–434 (1978).
- Eardley, D. D., Shen, F. W., Cantor, H. & Gershon, R. K. *J. exp. Med.* **150**, 44–50 (1979).
- Weinberger, J. Z., Greene, M. I., Benacerraf, B. & Dorf, M. E. *J. exp. Med.* **149**, 1336–1348 (1979).

Chronic immune stimulation is required for Moloney leukaemia virus-induced lymphomas

John C. Lee & James N. Ihle

Biological Carcinogenesis Program, Frederick Cancer Research Center, Frederick, Maryland 21701

C-type viruses are known to be aetiologically related to naturally occurring leukaemia in a variety of species¹, although the mechanisms of transformation are largely unknown. The long latency periods, requirement for an acute viraemia and monoclonality of the tumours^{2,3} distinguish leukaemogenesis from neoplasias induced by the acute transforming viruses and suggest an indirect mechanism. Consistent with this is the lack of *in vitro* transforming activity of replication-competent leukaemogenic viruses⁴. We have shown previously that the presence of T cells which proliferate *in vitro* in response to viral antigens is uniquely associated with the conditions leading to leukaemia. Based on these observations we have hypothesized that chronic immune stimulation is required for leukaemogenesis. We now demonstrate that CBA/N mice, when inoculated as newborns with Moloney leukaemia virus (MoLV), develop an acute viraemia but do not develop leukaemia or have detectable T-cell responses against the virus. This supports the hypothesis that chronic immune stimulation is essential for leukaemogenesis.

Previous studies have shown that MoLV can induce leukaemia in a variety of strains of mice⁵. However, as illustrated in Fig. 1, CBA/N mice failed to develop leukaemia after inoculation with a MoLV preparation which rapidly induced leukaemia in BALB/c mice. This result was surprising because the closely related CBA/J strain has been shown to be as susceptible to MoLV as BALB/c mice⁵. We therefore examined the virological and immunological status of MoLV-inoculated CBA/N mice.

The ability of the virus to establish viraemia is known to be essential for leukaemogenesis^{6,7}. To obtain a quantitative measure of virus replication we examined spleen extracts from inoculated mice for the viral proteins gp71, p30 and p12 using competition radioimmunoassays. As shown in Table 1, these proteins were readily detectable in CBA/N mice at 2 or 4 months post-inoculation. For comparison, the range of values obtained with MoLV-inoculated BALB/c or CBA/J mice are also shown. At 2 or 4 months post-inoculation, there were minimal differences in viral protein concentrations among the various strains. One important parameter which is known to be inversely related to leukaemia and viraemia is the presence of antibody against the virus⁸. Therefore, we examined the above

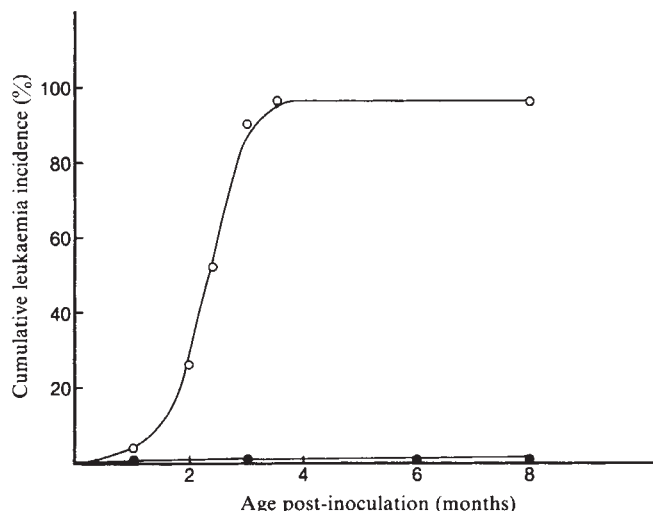


Fig. 1 Cumulative leukaemia incidence in MoLV-inoculated BALB/c or CBA/N mice. Groups of 20 BALB/c (○) or CBA/N (●) mice were inoculated as newborns with MoLV. The mice were reared in specific pathogen-free conditions. Moribund mice were killed and assessed histopathologically for the presence of lymphomas. The lymphomas in BALB/c mice were both of thymic and splenic origin, as previously described²³.

mice for the presence of free antibody against MoLV. As shown in Table 1 none of the mice examined had detectable antobodies. These data demonstrate that the absence of leukaemia in CBA/N mice is not due to an inability of MoLV to replicate and establish viraemia.

It is well established that CBA/N mice differ from CBA/J mice in an immune response function and fail to develop immune responses to certain antigens early in life⁸. This deficiency has been shown to be due to a X-linked recessive locus⁹. Because our previous results demonstrated that the presence of T cells which respond *in vitro* by proliferation against viral proteins is correlated with leukaemogenesis^{6,10}, we compared this cellular immune response of MoLV-inoculated BALB/c, CBA/J and CBA/N mice. Splenic lymphocytes were obtained from individual mice, partially purified by passage through nylon wool columns as previously described¹⁰ and assayed in a blastogenesis assay for proliferation in response to MoLV gp71. The responses of lymphocytes from individual BALB/c, CBA/J and CBA/N mice to various antigen concentrations are shown in Fig. 2. Lymphocytes from normal BALB/c, CBA/J and CBA/N did not respond to MoLV gp71. However, as previously demonstrated¹⁰ and shown in Fig. 2, lymphocytes from preleukaemic BALB/c and CBA/J mice

Table 1 Levels of MoLV proteins and antibody against MoLV in inoculated BALB/c, CBA/J and CBA/N mice

Strain	Age (months)	Protein (ng per mg)			Antibody titre
		gp71	p30	p12	
BALB/c	1	100–500	250–900	3–7	<1:20
BALB/c	2	300–1,400	400–1,100	10–50	<1:20
CBA/N	2	400–800	500–700	5–20	<1:20
BALB/c	3	900–1,700	700–1,500	60–70	<1:20
CBA/N	4	500–850	500–650	15–20	<1:20
CBA/J	4	1,500–1,900	1,000–1,200	30–50	<1:20

BALB/c, CBA/N and CBA/J mice were inoculated with MoLV at birth. At the indicated ages post-inoculation three to five individual mice were killed and used to determine the levels of antibody and viral proteins. The levels of viral proteins were determined using splenic extracts in competition radioimmunoassays for MoLV gp71, p30 and p12 as previously described²¹. The values represent the range of viral protein concentrations observed. The presence or absence of antibody in the serum was determined using ³H-leucine-labelled intact MoLV virions in a radioimmuno precipitation assay as previously described²².

responded *in vitro* to MoLV gp71 with typical antigen concentration dependence. Previous studies¹⁰ have shown this proliferative response to be dependent on a Thy-1.2⁺, Lyt-1⁺, 2⁻ T-cell subpopulation. In distinct contrast, however, lymphocytes from viraemic CBA/N mice had no detectable responses against MoLV gp71. Therefore, viraemic CBA/N mice seemed to be unique in lacking T cells specific for MoLV gp71.

We have also demonstrated that preleukaemic BALB/c mice have lymphocytes which respond *in vitro* to the internal viral protein p12 (ref. 10), although less consistently than to MoLV gp71. As shown in Fig. 2, lymphocytes from both preleukaemic BALB/c and CBA/J mice responded to p12 and showed typical antigen titration curves. Normal control BALB/c or CBA/J mice did not respond. The response to p12 in preleukaemic BALB/c and CBA/J mice was also found to depend on a Thy-1.2⁺, Lyt-1⁺, 2⁻ T cell (data not shown), whereas the MoLV-inoculated CBA/N mice had no detectable responses to p12.

We next examined individual MoLV-inoculated BALB/c, CBA/J and CBA/N mice at various ages. The results are summarized in Fig. 3, which presents the stimulation indices (SIs) against MoLV gp71 or p12. When lymphocytes from uninoculated BALB/c mice were examined, none had a significant blastogenic response (SI > 2.0) against gp71 or p12. In contrast, the majority of MoLV-inoculated BALB/c mice (86%) responded to gp71, whereas ~41% responded to p12.

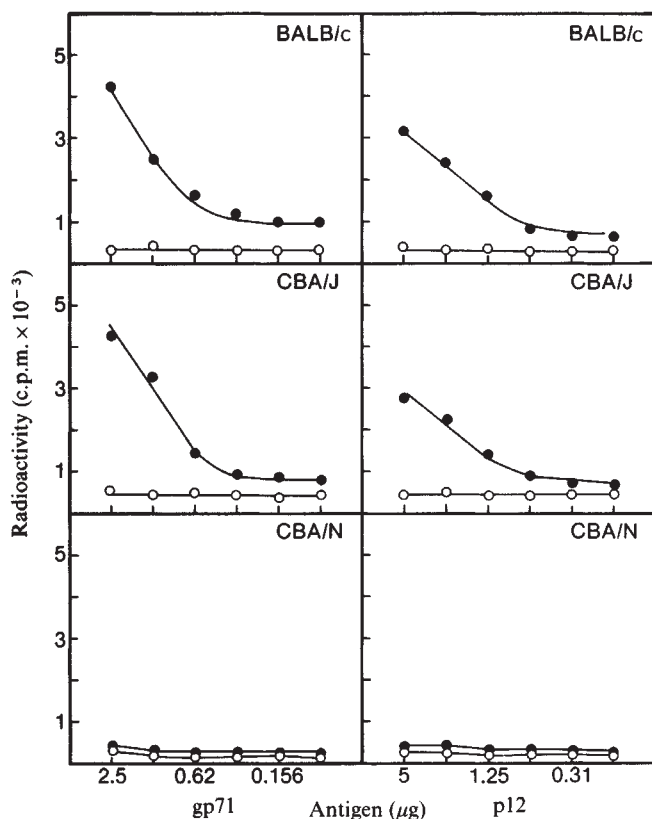


Fig. 2 T-cell proliferative responses of BALB/c, CBA/J and CBA/N mice to MoLV gp71 and p12. Lymphocytes were obtained from either 6-week-old normal mice (○) or mice inoculated as newborns with MoLV (●). An enriched T-cell preparation was obtained by nylon wool column fractionation and assayed in a microblastogenesis assay as previously described²⁴. The lymphocytes were stimulated *in vitro* using purified MoLV gp71 (left panels) or p12 (right panels) at the indicated concentrations. The purification techniques have been previously described²⁴ and the antigens used were homogeneous as assessed by SDS-polyacrylamide gel electrophoresis. After 3 days of incubation the cells were pulse-labelled with ³H-thymidine for 8 h and the incorporated radioactivity determined. Samples were done in triplicate and the individual samples varied less than 10% from the mean.

Interestingly, only one mouse (indicated by the solid circles in Fig. 3) failed to respond to either antigen. Using lymphocytes from individual MoLV-inoculated CBA/J mice, all the mice responded to MoLV gp71 *in vitro*, whereas five of the seven mice responded to MoLV p12. In distinct contrast, however, most of the CBA/N mice did not respond to either gp71 or p12. Only one mouse (indicated by the solid circles) had lymphocytes which responded at low levels to both MoLV gp71 and p12. These data demonstrate that MoLV-inoculated viraemic CBA/N mice are distinct from the other strains and generally lack a T-cell response against the viral proteins.

Our results describe two unique characteristics of the response of CBA/N mice to MoLV, that is, the lack of leukaemia and the lack of T cells which respond to MoLV proteins. The lack of leukaemia in viraemic mice is quite unusual. Gisselbrecht⁵ has reported that DBA/2 mice fail to develop leukaemia by 6 months of age after MoLV inoculation at 4 weeks of age. However, we have observed that inoculated newborn DBA/2 mice develop leukaemia with a cumulative incidence of 85% by 8 months after inoculation (unpublished data). Interestingly, of the DBA/2 mice examined in those studies all had T-cell proliferative responses to MoLV gp71. Therefore, we believe that CBA/N represents the only mouse strain in which MoLV-induced viraemia is not associated with leukaemia.

Although several genetic factors are known to influence MoLV-induced leukaemogenesis⁵, most of these factors phenotypically decrease the level of viraemia. Our results clearly demonstrate that there were no significant differences in viraemia between the susceptible BALB/c and CBA/J strains and the resistant CBA/N mice. Another possible reason for the absence of leukaemia is the lack of an appropriate recombinant oncogenic virus. It has been suggested that spontaneous leukaemias in AKR mice require recombination to generate oncogenic MCF viruses^{11,12}. In MoLV-inoculated CBA/N mice this seems unlikely for two reasons. First, cloned MoLV is oncogenic in rats and only the parental virus is found in most lymphomas (ref. 2 and D. Steffen, personal communication); therefore, recombination does not seem to be required. Second, the stock of MoLV used in these studies is an uncloned animal passaged stock which contains recombinant viruses similar to those previously described¹³. Nevertheless, the oncogenic potential and host range characteristics of viruses recovered from CBA/N mice are now being evaluated.

Our results demonstrate that viraemic CBA/N mice lack lymphocytes that respond to MoLV proteins whereas such lymphocytes are found in viraemic mice of other strains^{6,10}. CBA/N mice are known to differ from CBA/J mice in their ability to respond to certain antigens⁸. However, the immune deficiency of CBA/N mice has only been shown to involve thymus-independent antigens, whereas our data suggest the absence of a T-cell function. One important distinction in our studies, however, is that MoLV was given to newborn mice, whereas other studies have examined the effect of immunizing adult mice. Therefore, the lack of responses may be related to the X-linked immune response gene and is manifested as tolerance. Preliminary genetic experiments support this in that the lack of a response to viral antigens is X linked.

The essential question is whether the absence of a detectable T-cell response against MoLV is responsible for the lack of leukaemia in CBA/N mice. Although the present results suggest this, appropriate genetic crosses will be required to confirm the relationship. If the lack of an immune response is involved, several possible explanations exist. McGrath and Weissman¹⁴ presented evidence that T cells with appropriate virus-specific receptors were required for leukaemogenesis. If these receptors represent the immune specific receptors of T cells, our results demonstrate that CBA/N mice lack this subpopulation of T cells. Second, it is conceivable that an immune response is required to generate appropriate target-cell populations for virus infection and/or transformation. In a somewhat analogous manner we have previously hypothesized that chronic immune stimulation is required for leukaemogenesis. The basic

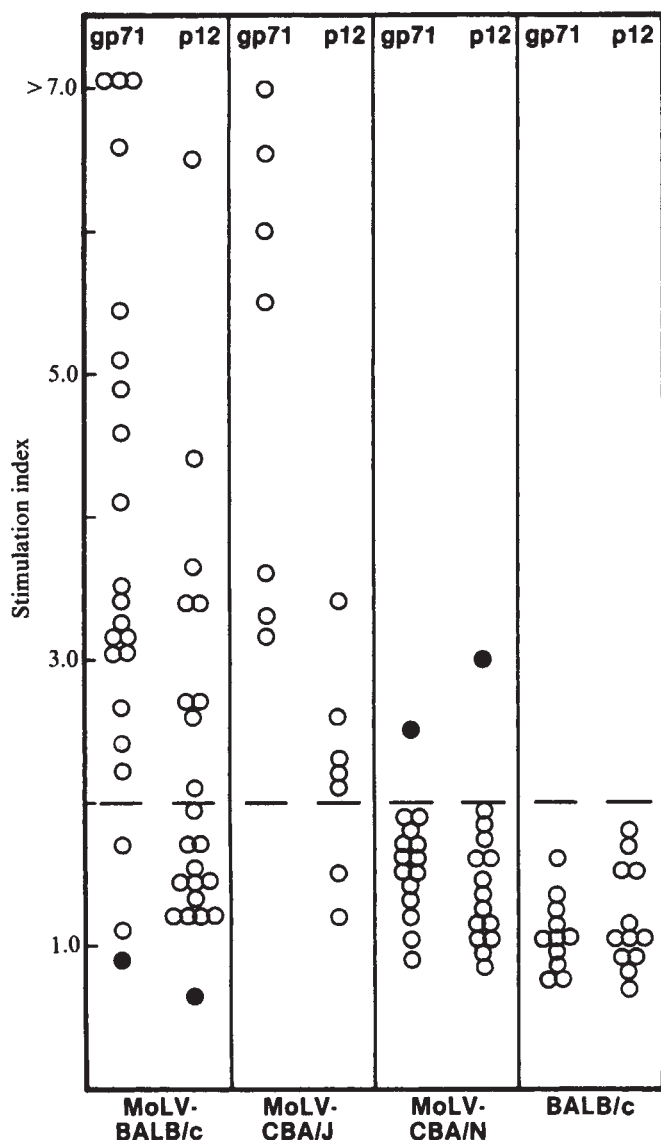


Fig. 3 T-cell proliferative responses of individual MoLV-inoculated or control BALB/c, CBA/N or CBA/J mice. Mice inoculated as newborns with MoLV or normal uninoculated BALB/c mice were used. Individual mice at 1–4 months of age were killed and splenic lymphocytes obtained and assayed in blastogenesis assays as described in Fig. 2 legend. Assays included the responses to MoLV gp71 (2.5 µg per well) or p12 (5 µg per well). All samples were tested in triplicate and the mean c.p.m. incorporated was determined. Individual samples generally varied less than 10% from the mean. The stimulation index presented is the ratio of c.p.m. of ^3H -thymidine incorporated in the presence of antigen to that in the absence of antigen. A stimulation index of 2.0 is considered a significant response. The solid circles indicate individual mice which either failed to respond to either antigen (BALB/c) or responded to both antigens (CBA/N), as discussed in the text.

components of this hypothesis are: (1) As a consequence of the establishment of an acute viraemia and expression of viral proteins, antigen reactive Thy 1.2⁺, Lyt-1⁺, 2⁻ T cells are continually recruited. (2) In response to antigen these T cells produce a variety of soluble lymphokines which influence the proliferation and differentiation of the immune system. Such factors include macrophage activating factors¹⁵ which in turn cause the production of lymphocyte activating factor (LAF) or interleukin I (ref. 16). In addition, blastogenic factors are produced which promote proliferation of early T cells¹⁷. Other factors include those related to T-cell growth factors which promote thymocyte differentiation¹⁸ and induction of differentiation of mature T cells¹⁹. (3) This chronic state of proliferation

and differentiation increases the probability of spontaneous events leading to transformation. One likely 'error' commonly found associated with murine leukaemias that may be associated with transformation would be the generation of a trisomy in chromosome 15 (ref. 20). Irrespective of the validity of this interpretation, the results described here should provide important approaches to studying the mechanisms by which C-type viruses induce leukaemia.

We thank Jonathan Keller, Kathleen Bengali and Linda Rebar for technical assistance and Linda Fawley for help in preparation of the manuscript. This work was supported by NCI contract no. NO1-CO-75380, NIH.

Received 8 August; accepted 17 November 1980.

1. Tooze, J. *The Molecular Biology of Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1973).
2. Steffen, D. & Weinberg, R. A. *Cell* **15**, 1003 (1978).
3. Canaani, E. & Aaronson, S. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1677 (1979).
4. Ihle, J. N. *et al.* in *Cold Spring Harb. Conf. Cell Proliferation* Vol. 7, 1049 (1980).
5. Gisselbrecht, S. *et al.* *Int. J. Cancer* **21**, 626 (1978).
6. Lee, J. C. & Ihle, J. N. *J. Immun.* **123**, 2351 (1979).
7. Lilly, F., Duran-Reynals, M. L. & Rowe, W. P. *J. exp. Med.* **141**, 882 (1975).
8. Amsbaugh, D. F., Hansen, C. T., Prescott, B., Barthold, D. R. & Baker, P. J. *J. exp. Med.* **136**, 931 (1972).
9. Berning, A. K., Eicher, E. M., Paul, W. E. & Scher, I. *J. Immun.* **124**, 1875 (1980).
10. Lee, J. C., Horak, I. & Ihle, J. N. *J. Immun.* (in the press).
11. Hartley, J. W., Wolford, N. K., Old, L. J. & Rowe, W. P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 789 (1977).
12. Cloyd, M. W., Hartley, J. W. & Rowe, W. P. *J. exp. Med.* **151**, 542 (1980).
13. Fischinger, P. J., Nomura, S. & Bolognesi, D. P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 5150 (1975).
14. McGrath, M. S. & Weissman, I. L. *Cell* **17**, 65 (1979).
15. Ng, A. K., Ames, R. S., McIntire, R. K. & Herberman, R. B. *Cancer Res.* **39**, 4887 (1979).
16. Farrar, J. J. & Koopman, W. J. in *Biology of the Lymphokines* (eds Cohen, S., Pick, E. & Oppenheim, J. J.) 325 (1979).
17. Enjuanes, L., Lee, J. & Ihle, J. N. *J. Immun.* (submitted).
18. Watson, J., Gillis, S., Marbrook, J., Mochizuki, D. & Smith, K. A. *J. exp. Med.* **150**, 849 (1979).
19. Ihle, J. N., Pepersack, L. & Rebar, L. *J. exp. Med.* (submitted).
20. Klein, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2442 (1979).
21. Fischinger, P. J. *et al.* *Virology* **90**, 241 (1978).
22. Ihle, J. N. & Joseph, D. R. *Virology* **87**, 287 (1978).
23. Pepersack, L., Lee, J. C., McEwan, R. & Ihle, J. N. *J. Immun.* **124**, 279 (1980).
24. Enjuanes, L., Lee, J. C. & Ihle, J. N. *J. Immun.* **122**, 665 (1979).

Phagocytosis requires repeated triggering of macrophage phagocytic receptors during particle ingestion

Denise R. Shaw & Frank M. Griffin Jr

Departments of Microbiology and Medicine, University of Alabama in Birmingham, Birmingham, Alabama 35294

The phagocytic functions of polymorphonuclear leukocytes and mononuclear phagocytes are crucial elements in host defence against a variety of invading microorganisms. Phagocytosis is a highly selective process, requiring specific interactions between the surface of the particle to be ingested and the plasma membrane of the phagocytic cell. The phagocyte can therefore discriminate between 'ingestible' and 'non-ingestible' particles even when the different particles are located in close proximity on the plasma membrane of the phagocytic cell¹. Previous work has demonstrated that these specific interactions between particle and phagocyte are required not only for the initiation of phagocytosis—that is, attachment of the particle to the phagocytic cells and generation of a signal to trigger phagocyte pseudopod extension—but also for the subsequent progression of pseudopods over the entire surface of the particle^{2,3}. We present evidence here that the continued interactions between phagocyte plasma membrane receptors and particle-bound ligands do not function merely to direct otherwise random phagocyte pseudopod movement, but instead are required for the repeated generation of intracellular phagocytic signals during the entire ingestion process.

All plasma membrane receptors for immunological ligands promote the binding of ligand-coated particles to the phagocytic cell surface; however, only certain receptors are able to generate a phagocytic signal, and this ability depends on the physiological state of the cell. Mouse peritoneal macrophages possess well characterized plasma membrane receptors of both phagocytic signal generating and non-generating classes. Receptors for the Fc portion of immunoglobulin G (IgG) molecules mediate both the attachment and the ingestion of particles coated with IgG (refs 4, 5). Receptors for the third component of complement (C3) mediate the attachment of particles coated with C3, but resident macrophages (R-MΦ) do not ingest the attached C3-coated particles⁵⁻⁷; however, inflammatory macrophages, such as those elicited by the intraperitoneal injection of Brewer's thioglycollate medium (TG-MΦ)⁸, and resident macrophages treated *in vitro* with certain lymphokine preparations (LT-MΦ)⁹ attach and ingest C3-coated particles.

Previous studies have indicated that the density of the immunological ligands on the surface of a particle can determine whether or not the particle will be phagocytosed. For example, particles coated with diminishing quantities of IgG are ingested by macrophages with progressively reduced efficiency; however, particles coated with inadequate amounts of IgG can be converted into particles which are readily ingested by R-MΦ after the addition of C3 ligands to the particle surface¹⁰. Other work has shown that if the phagocytosis-triggering ligands on an ingestible particle are sequestered on one hemisphere of the particle surface, that particle cannot be ingested³. From these previous experimental observations, we considered that there were two possible hypotheses for the role of phagocyte plasma membrane receptors in the process of particle ingestion: (1) Once a phagocytic cell has been triggered by a certain minimum number of phagocytic signals generated by appropriate ligand-plasma membrane receptor interactions, the phagocyte is committed to pseudopod extension of a degree sufficient to result in interiorization of the bound particle. Subsequent ligand-receptor engagement is required only to guide this pseudopod movement towards particle engulfment. Without such direction, fusion of the leading edges of pseudopods around the attached particle would be a random, inefficient process. (2) The engagement of phagocytic signal-generating plasma membrane receptors by the corresponding ligands triggers the phagocytic cell to produce only a limited amount of pseudopod extension. Unless such movement results in additional engagement of signal-generating receptors, particle engulfment will cease. Ligand-receptor interactions which do not generate phagocytic signals may function to facilitate each increment of

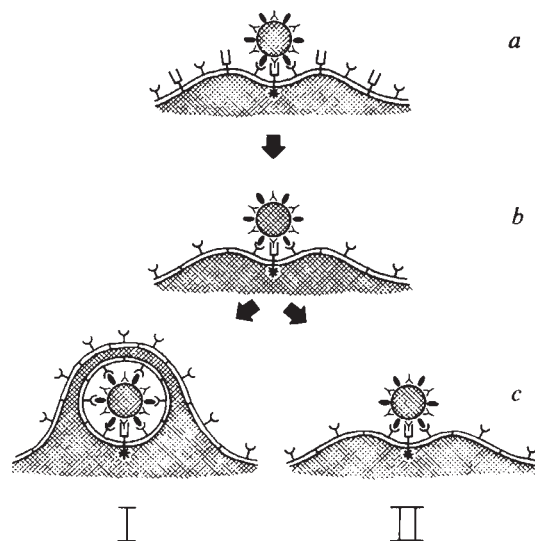


Fig. 1 Schematic representation of the experiment. *a*, IgG + C3-coated erythrocytes were bound to macrophage plasma membranes at 4 °C through the interactions of particle-bound IgG (Y) and C3 (●) ligands with the macrophages' Fc (Y) and C3 receptors (●), respectively. *b*, Fc receptors located on the remainder of the macrophages' surfaces were functionally removed by treatment with anti-macrophage serum at 4 °C. *c*, Macrophages were then incubated at 37 °C for 30 min and examined for erythrocyte ingestion. If the initial phagocytic signal-generating (*) ligand-receptor (IgG-Fc) engagement triggers sufficient pseudopod extension, R-MΦ should complete erythrocyte ingestion through the 'passive' binding of C3 receptors to C3 ligands on the remainder of the erythrocyte surface (result I). However, if repeated phagocytic signal-generating ligand-receptor interactions are required for each increment of pseudopod extension, then IgG + C3-coated erythrocytes should not be ingested by R-MΦ (result II).

pseudopod movement over the surface of the particle and thereby increase the efficiency of signal-generating receptor binding to additional ligands on the particle, but they cannot by themselves complete the process of particle engulfment.

To discern which of the two hypotheses best describes the mechanism of phagocytosis, we have performed the experiment schematically shown in Fig. 1. First, macrophages were allowed to attach, through their Fc and C3 receptors, particles coated with both IgG and C3, in conditions which prevented particle

Table 1 Attachment and ingestion of erythrocytes by macrophages

Erythrocytes coated with:	Attachment index $\pm$ s.e.m.			Phagocytic index $\pm$ s.e.m.		
	R-MΦ	TG-MΦ	LT-MΦ	R-MΦ	TG-MΦ	LT-MΦ
Control MΦ						
1. Nothing	8.1 $\pm$ 1.6 (3)	20 $\pm$ 9.1 (3)	40 $\pm$ 17 (3)	0 (3)	7.4 $\pm$ 4.7 (5)	7.9 $\pm$ 6.4 (3)
2. C3	1,200 $\pm$ 100 (3)	940 $\pm$ 28 (2)	840 $\pm$ 140 (3)	9.8 $\pm$ 6.8 (3)	490 $\pm$ 180 (2)	350 $\pm$ 85 (3)
3. IgG	510 $\pm$ 139 (4)	1,100 $\pm$ 290 (4)	1,100 $\pm$ 200 (3)	300 $\pm$ 59 (5)	700 $\pm$ 230 (4)	670 $\pm$ 96 (4)
4. IgG + C3	710 $\pm$ 280 (4)	970 $\pm$ 160 (4)	2,000 $\pm$ 310 (3)	440 $\pm$ 110 (5)	610 $\pm$ 140 (4)	1,000 $\pm$ 54 (4)
MΦ treated with anti-macrophage serum						
5. IgG	520 $\pm$ 100 (7)	960 $\pm$ 250 (4)	1,200 $\pm$ 280 (4)	13 $\pm$ 4.2 (8)	29 $\pm$ 9.5 (4)	53 $\pm$ 6.5 (8)
6. IgG + C3	750 $\pm$ 130 (7)	900 $\pm$ 320 (5)	1,700 $\pm$ 170 (4)	19 $\pm$ 9.5 (8)	460 $\pm$ 48 (5)	310 $\pm$ 24 (8)

Macrophages from non-treated mice (R-MΦ) or from thioglycollate-injected mice (TG-MΦ) were cultured for 48 h before the assay. Some macrophages from non-treated mice were cultured in medium containing the previously described lymphokine (LT-MΦ) which enhances macrophage C3 receptor function⁹. After attachment of erythrocytes to macrophages, some cultures were treated with anti-macrophage serum at a concentration that reduced Fc receptor-mediated phagocytosis by 90% but was not toxic to macrophages. Results are expressed as attachment or phagocytic indices, calculated by multiplying the percentage of macrophages attaching or ingesting one or more erythrocytes by the mean number of erythrocytes attached or ingested per macrophage. Numbers in parentheses indicate the number of determinations, each determination being made on 100 macrophages. The macrophage cultures were relatively homogeneous, and attachment and phagocytic indices are representative of the population as a whole. Thus, small indices (<60) were obtained only when less than 20% of macrophages attached or ingested a few erythrocytes each, whereas large indices (>300) were never the result of a minority of macrophages attaching or ingesting large numbers of erythrocytes but were obtained only when greater than 65% of the population participated in the function assessed.

ingestion (Fig. 1a). Non-engaged macrophage Fc receptors were specifically blocked without affecting the preformed receptor interactions (Fig. 1b). Cultures were then restored to conditions permissive for ingestion (Fig. 1c). The protocol used was intended to provide an initial phagocytosis trigger in the form of the engagement of macrophage Fc receptors by particle-bound IgG. However, completion of particle ingestion, if it occurred, would have to proceed through macrophage C3 receptor binding to particle-bound C3, interactions which alone are unable to promote ingestion by R-M Φ .

Procedures for obtaining monolayer cultures of R-, TG- and LT-M Φ and for preparing sheep erythrocytes (particles) coated with IgG, C3 or IgG+C3 have been described in detail elsewhere^{1-3,7-9}. Erythrocytes were incubated with macrophage cultures at 4 °C for 30 min, conditions which allow the attachment but not the ingestion of ligand-coated particles¹¹. Cultures were then treated for 60 min at 4 °C with the IgG fraction of a rabbit anti-mouse macrophage serum which has previously been shown to block specifically macrophage Fc receptor function without affecting macrophage C3 receptors^{8,12}. Cultures were then transferred to 37 °C and an atmosphere of 95% air-5% CO₂, conditions which are permissive both for phagocytic signal transmission by IgG-Fc receptor bonds previously formed at 4 °C (ref. 3) and for particle ingestion^{2,3}. After incubation for 30 min, cultures were prepared for phase contrast microscopy and scored for erythrocyte phagocytosis as previously described¹.

Control experiments (Table 1, lines 1-5) confirmed that non-coated erythrocytes were neither bound nor ingested by any macrophage preparation (line 1), that C3-coated erythrocytes were attached but not ingested by R-M Φ and were ingested by TG- and LT-M Φ (line 2), that erythrocytes coated with IgG (line 3) and those coated with IgG+C3 (line 4) were attached and ingested by all three types of macrophages, and that treatment with anti-macrophage serum did not affect the previously established attachment of IgG-coated erythrocytes, but did efficiently block the subsequent ingestion of these erythrocytes by all macrophage preparations (compare lines 3 and 5).

As expected, TG- and LT-M Φ were able to ingest IgG+C3-coated erythrocytes even after treatment with anti-macrophage serum (line 6, columns 5, 6). The phagocytic indices were nearly identical to those obtained when untreated TG- and LT-M Φ were incubated with erythrocytes coated with C3 alone (line 2), in agreement with previous observations that anti-macrophage serum treatment completely blocks macrophage Fc receptors but does not impair macrophage C3 receptor function^{8,12}. In contrast, R-M Φ treated with anti-macrophage serum were unable to ingest IgG+C3-coated erythrocytes (line 6, column 4).

The inability of R-M Φ to ingest IgG+C3-coated erythrocytes supports hypothesis (2). These erythrocytes presented sufficient C3 ligands so that, even after macrophage Fc receptors had been blocked, they could be ingested by TG- and LT-M Φ , macrophages for which C3 receptor engagement is capable of generating phagocytic signals. However, in similar conditions, R-M Φ were unable to ingest these IgG+C3-coated particles, even though the initial interaction of particle with macrophage could occur by phagocytic signal-generating ligand-receptor (IgG-Fc) engagements and even though the R-M Φ could efficiently attach particles through their C3 receptors.

We thank Ms Peggy M. Rains for technical assistance and Dr Roy Curtiss III for very helpful criticism of the manuscript. This work was supported by American Cancer Society grant IM-173. D.R.S. is supported by NIH Predoctoral Training grant GM-07164. F.M.G. is the recipient of NIH Research Career Development Award AI-00135.

Received 22 August; accepted 13 November 1980.

- Griffin, F. M. & Silverstein, S. C. *J. exp. Med.* **139**, 323-336 (1974).
- Griffin, F. M., Griffin, J. A., Leider, J. E. & Silverstein, S. C. *J. exp. Med.* **142**, 1263-1282 (1975).
- Griffin, F. M., Griffin, J. A. & Silverstein, S. C. *J. exp. Med.* **144**, 788-809 (1976).

- Berken, A. & Benacerraf, B. *J. exp. Med.* **123**, 119-144 (1966).
- Mantovani, B., Rabinovitch, M. & Nussenzweig, V. *J. exp. Med.* **135**, 780-792 (1972).
- Lay, W. H. & Nussenzweig, V. *J. exp. Med.* **128**, 991-1009 (1968).
- Griffin, F. M., Bianco, C. & Silverstein, S. C. *J. exp. Med.* **141**, 1269-1277 (1975).
- Bianco, C., Griffin, F. M. & Silverstein, S. C. *J. exp. Med.* **141**, 1278-1290 (1975).
- Griffin, J. A. & Griffin, F. M. *J. exp. Med.* **150**, 653-675 (1979).
- Ehlenberger, A. G. & Nussenzweig, V. *J. exp. Med.* **145**, 357-371 (1977).
- Rabinovitch, M. *Exptl Cell Res.* **46**, 19-28 (1967).
- Holland, P., Holland, N. H. & Cohn, Z. A. *J. exp. Med.* **135**, 458-475 (1972).

Nonjunctional acetylcholine receptor channel open time decreases during development of *Xenopus* muscle

R.W. Kullberg

Biology Department, University of Alaska, Anchorage, Anchorage, Alaska 99504

P. Brehm & J. H. Steinbach

The Salk Institute, PO Box 85800, San Diego, California 92138

The mean channel open time of junctional acetylcholine receptors (AChRs) decreases during development of mammalian skeletal neuromuscular junctions¹⁻³. The junctional AChRs of newborn rats have long mean channel open times of about 4.5 ms and in this respect are similar to nonjunctional AChRs of adult denervated muscle³⁻⁵. Mature neuromuscular junctions have channels with shorter mean open times of about 1 ms, whereas junctions at intermediate stages seem to have a mixture of channels with short and long open times¹⁻³. These observations, taken together with findings on receptor distribution during development^{6,14}, suggest that the establishment of nerve-muscle contact induces an aggregation of embryonic AChRs with long mean channel open times at the point of contact and that these embryonic junctional receptors are then replaced, modified or subjected to local membrane changes leading to a discrete reduction in channel open time. However, the fate of those embryonic nonjunctional receptors which are not aggregated at the newly formed neuromuscular junction is unknown. Are they an unchanging population, or do the changes in kinetic behaviour which occur at the junctional regions also take place at other regions of the muscle membrane? The kinetics of nonjunctional receptors in mammalian skeletal muscle have not been studied during development. We have now, however, examined this question in amphibian myotomal muscle, where nonjunctional receptors are present in substantial numbers throughout development, making it possible to measure the mean open time of both classes of AChR channel. We observed a shortening of mean channel open time of junctional AChRs, from about 3 ms to 1 ms, and an identical change in mean channel open time of AChRs at nonjunctional regions of membrane.

The experiments were performed on myotomal muscles of *Xenopus laevis* embryos and tadpoles. Procedures of dissection and staging have been described previously⁷. The muscle cells extend the length of a myotome and insert at myocommata where the neuromuscular junctions are located. Focal synaptic potentials may be recorded with an extracellular pipette positioned at the myocommata⁸, and staining with fluorescent-labelled α -bungarotoxin reveals extensive receptor accumulations at the ends of the muscles⁹. Muscle cells ranged in length from 100 μ m in the earliest stages studied to 200 μ m in the latest stages.

Recordings of ACh-induced extracellular noise were obtained from muscles using the method developed by Schuetze, Frank and Fischbach¹⁰. An extracellular pipette (3-7 μ m o.d.) filled with bathing solution (see Fig. 1 legend) containing either 25 or 50 μ M ACh was pressed against the muscle to record agonist-induced noise. Recordings were

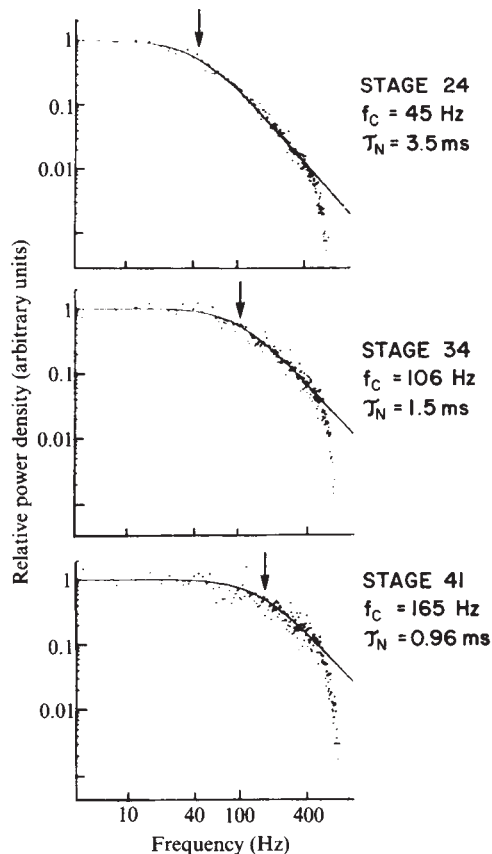


Fig. 1 Sample spectra for fluctuations recorded from non-junctional regions of *Xenopus* myotomal muscle cells at three different developmental stages. The spectra are plotted logarithmically and the power is expressed in arbitrary units. Data were bandpass filtered at 0.5–500 Hz and digitized at 0.49-ms intervals. Spectra were computed with a fast Fourier-transform routine using a Digital PDP 11/34 computer. The spectra shown have had background spectra subtracted (this correction resulted in a change of less than 10% in the power at frequencies less than approximately twice the half-power frequency). The background spectra were obtained from records in which the pipette was not pressed against a cell. Single Lorentzian curves are shown, fitted to the data by eye. A Lorentzian curve is described by the equation $S(f) = S_0/[1 + (f/f_c)^2]$, where S_0 is the zero frequency power, f_c is the half-power frequency and f is frequency. The recording solution contained 111 mM NaCl, 3 mM KCl, 1.8 mM CaCl₂ and 8 mM HEPES-NaOH buffer (pH 7.4). Records were made at room temperature (20–23 °C).

classified as 'nonjunctional' when the electrode tip was placed at the middle of the myotome, where no synapses are located, and 'junctional' when positioned at the myocommata. In the latter case, the occurrence of infrequent, miniature endplate currents (m.e.p.c.s) in some recordings confirmed the close proximity of the electrode tip to junctional regions. To avoid possible differences in myotomal development along the length of the tadpole, only the anterior myotomes were included in this study. Sealing of the ACh-containing pipette against the muscle membrane (sealing resistance 0.5–5 MΩ) was accompanied by an increase in voltage fluctuations and a mean negative voltage shift. This signal resulted from the activity of ACh-activated channels, as it was abolished by (+) tubocurarine (3×10^{-5} M) and was not observed when the electrode was either lacking ACh or pressed against epidermis.

Estimates of mean channel open time were obtained from spectral analysis of ACh-induced noise. Power density spectra were computed and fitted by single Lorentzian curves (Fig. 1) which describe the theoretical spectra produced by randomly

opening ACh-activated channels¹¹. In conditions of low ACh concentration, the half-power frequency (f_c) of a spectrum of fluctuations recorded from receptors on adult amphibian muscle is determined only by the channel-closing rate constant and provides an estimate (τ_n) of the mean channel open time according to the equation $\tau_n = 1/(2\pi f_c)$ (ref. 11). At high concentrations of ACh, τ_n will underestimate the true mean channel open time, because the half-power frequency is then determined by both channel opening and closing rates^{11,12}. To minimize the error in our estimate of mean channel open time, we have used the lowest possible ACh concentration which gave consistently satisfactory recordings. Because of dilution, the ACh concentration at the tip of the recording pipette was less than 25–50 μM. In most cases, a pipette was used until dilution rendered the ACh concentration so low that no further response could be recorded.

Spectral analysis of both junctional and nonjunctional AChR channel fluctuations was performed at developmental stages between 2 and 58 h after initial nerve contact. The power in the fluctuations extended to higher frequencies at later stages, yielding estimates of mean channel open times of the AChRs that are shorter in later developmental stages (Fig. 1). Values of τ_n were determined from spectra recorded at 101 sites (39

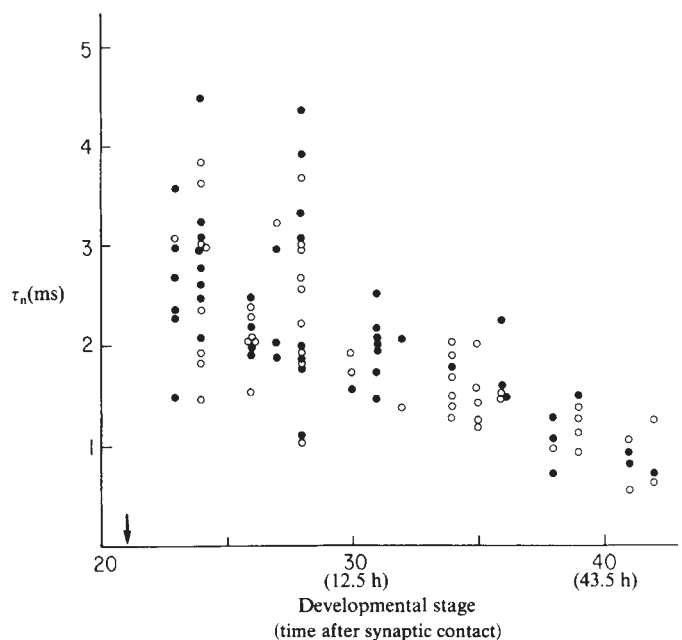
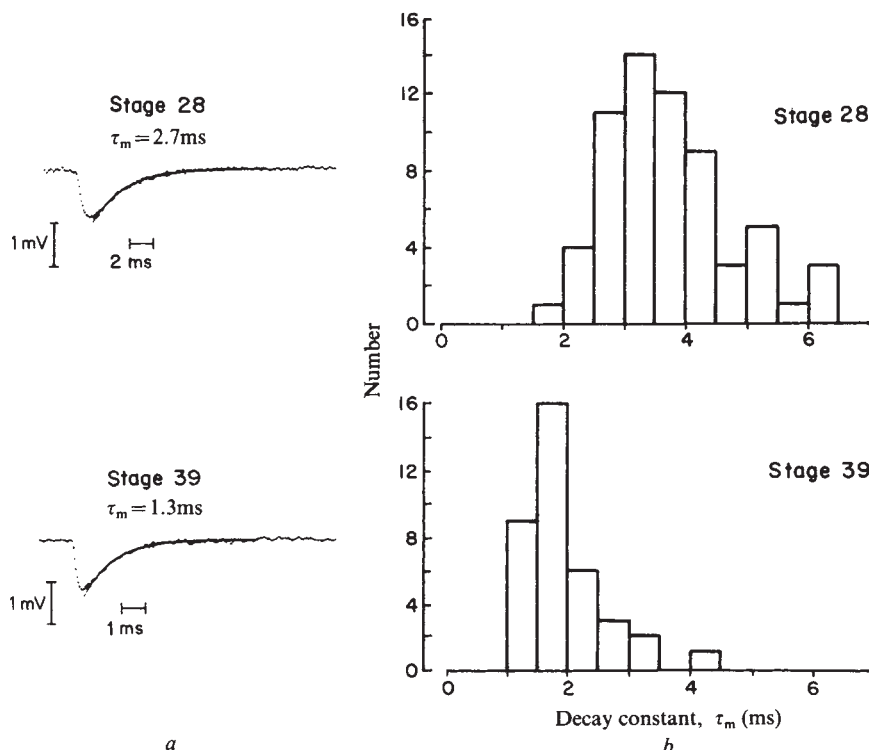


Fig. 2 The calculated τ_n for junctional (filled circles) and non-junctional (open circles) AChRs are plotted as a function of developmental stage. The abscissa is also marked in terms of hours of development after the first signs of synaptic transmission^{7,15}. The arrow indicates the time when the first synaptic activity may be recorded. Each point gives τ_n for a single recording site.

embryos). The values of τ_n measured at junctional regions decreased from 3.0 ± 0.7 ms (mean $\pm$ s.d., $n = 8$) at stage 24 to 1.1 ± 0.3 ms ($n = 7$) at stages 38–42. These values were not significantly different from the corresponding values for the nonjunctional AChRs, 2.7 ± 0.9 ms ($n = 8$) at stage 24 and 1.2 ± 0.3 ms ($n = 9$) at stages 38–42 (Fig. 2).

Developmental changes in junctional AChR channel kinetics were also examined by analysis of focally recorded synaptic currents (focal m.e.p.c.s). In conditions where the transient rise and fall of cleft concentration of transmitter is rapid relative to channel closing kinetics, the decay time constant of synaptic currents provides an independent estimate (τ_m) of the mean

Fig. 3 *a*, Focal m.e.p.c.s recorded at stages 28 and 39. Recordings were made by positioning the tip of a Ringer-filled extracellular pipette at the myocommata. Focal m.e.p.c.s were bandpass filtered (0.5–2 kHz) and sampled on-line by computer. The individual m.e.p.c. decays were then fitted by the best least-squares exponential curve. These curves are shown superimposed on the m.e.p.c. decay phases. *b*, Histograms of decay constants of focal m.e.p.c.s at stages 28 and 39. Each histogram is based on a sample of focal m.e.p.c.s recorded at a single site.



channel open time^{11,13}. In developing rat neuromuscular junctions, there is a fairly good correspondence between the time constant of m.e.p.c. decay and spectral estimates of mean channel open time^{1–3}. We recorded samples of focal m.e.p.c.s at selected stages of development, and the decay phases were fitted by single exponential curves (Fig. 3*a*). At any single recording site, there was a wide scatter of estimated decay constants, many of which were longer than the spectral estimates of mean channel open time at the same stage (Fig. 3*b*). However, the shorter decay constants, as illustrated by the m.e.p.c.s in Fig. 3*a*, were similar to spectral estimates of the mean channel open time at each stage. We never observed focal m.e.p.c.s which had briefer time constants than the fastest spectral estimates of mean channel open time at the same stage. These facts suggest that in the developing myotomal neuromuscular junction there is a fraction of ACh release sites where the decay of synaptic current is limited by channel kinetics and not by the rate of disappearance of transmitter.

We were unable to obtain measurements of τ_n in animals beyond stage 42 using the focal noise technique, possibly because of a decrease in nonjunctional receptor density and an increase in connective tissue at the myocommata, which made the receptors less accessible to the ACh-containing electrode. We were, however, able to estimate τ_m from focal m.e.p.c.s. In stage 50 muscle, τ_m was 0.7 ± 0.1 (mean $\pm$ s.d., $n = 181$ focal m.e.p.c.s; 6 recording sites); this indicates that receptor kinetics continue to accelerate beyond our stage 42 noise measurements.

The acceleration of junctional AChR channel kinetics observed at the mammalian neuromuscular junction results from a discrete shift to shorter mean channel open times^{1–3}. If this mechanism explained the shift in τ_n seen for *Xenopus* AChRs, the spectra at intermediate stages should be fitted by the sum of two lorentzian functions corresponding to the two populations of receptors. The values for τ_n would be expected to correspond, as in the case of rat muscle, to the slowest and fastest estimated values of mean channel open time at early and late stages. However, of 69 spectra recorded from intermediate stages (26–39), 50 spectra were satisfactorily fitted by eye with single lorentzian curves, as in Fig. 1, stage 34. The remaining spectra seemed to be fitted better with double lorentzian curves

composed of slow and fast components. Further examination of high resolution spectra will be necessary to determine whether the changes in channel kinetics occur gradually or discretely.

The importance of the present study is the finding that a modification of AChR channel kinetics occurs in both junctional and nonjunctional AChRs in developing *Xenopus* myotomal muscle. The process of change in AChR channel kinetics does not seem to depend on the proximity of receptors to the site of nerve contact with muscle, but affects receptors over the entire surface membrane. The development of the mammalian skeletal neuromuscular junction is similar to that of the myotomal synapse in that nonjunctional receptors are present in significant numbers ($150 \mu\text{m}^{-2}$)¹⁴ during the period of transition from slow to fast kinetics of junctional receptors. Given the present observations of *Xenopus* embryonic receptors, it would be interesting to see if similar changes in the kinetics of mammalian embryonic nonjunctional receptors occur during neuromuscular junction development.

This work was supported by NIH and NSF grants to J.H.S. and a Muscular Dystrophy Association grant to R.W.K. P.B. was supported by a Muscular Dystrophy Association Postdoctoral Fellowship. We thank Drs Y. Kidokoro, S. Halegoua, V. Dionne and R. Jaffe for critical reading of this paper.

Received 11 August; accepted 15 October 1980.

1. Sakmann, B. & Brenner, H. R. *Nature* **276**, 401–402 (1978).
2. Schuetze, S. M. & Fischbach, G. D. *Soc. Neurosci. Abstr.*, 8th a. Meet. 374 (1978).
3. Fischbach, G. D. & Schuetze, S. M. *J. Physiol., Lond.* **303**, 125–137 (1980).
4. Dreyer, F., Muller, K.-D., Peper, K. & Sterz, R. *Pflügers Arch. ges. Physiol.* **367**, 115–122 (1976).
5. Sakmann, B. *Fedn Proc.* **37**, 2654–2659 (1978).
6. Anderson, M. J. & Cohen, M. W. *J. Physiol., Lond.* **268**, 757–773 (1977).
7. Kullberg, R. W., Cohen, M. W. & Lentz, T. L. *Dev. Biol.* **60**, 101–129 (1977).
8. Kullberg, R. W., Mikelberg, F. S. & Cohen, M. W. *Dev. Biol.* **75**, 255–267 (1980).
9. Anderson, M. J. & Cohen, M. W. *J. Physiol., Lond.* **237**, 385–400 (1974).
10. Schuetze, S. M., Frank, E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 520–523 (1978).
11. Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–691 (1973).
12. Sakmann, B. & Adams, P. R. *Proc. 7th int. Congr. Pharmac.*, Paris, 1978 (ed. Jacob, J.) 81–90 (Pergamon, Oxford, 1979).
13. Colquhoun, D., Large, W. A. & Rang, H. P. *J. Physiol., Lond.* **266**, 361–395 (1977).
14. Bevan, S. & Steinbach, J. H. *J. Physiol., Lond.* **267**, 195–213 (1977).
15. Blackshaw, S. & Warner, A. *Nature* **262**, 217–218 (1976).

Interferon action—sequence specificity of the ppp(A2'p)_nA-dependent ribonuclease

Daniel H. Wreschner*, John W. McCauley*,
John J. Skehel & Ian M. Kerr*

National Institute for Medical Research, Mill Hill,
London NW7 1AA, UK

The oligonucleotides pppA2'p5'A2'p5'A and related oligomers (2-5A) are synthesized by an enzyme that is widely distributed in a variety of cells, the activity of which varies with interferon treatment, growth and hormone status^{1,2}. Because significant amounts of 2-5A have recently been detected in interferon-treated cells³, it has been suggested that the oligonucleotides may be involved in interferon action and in the control of cell metabolism. In both intact cells^{4,5} and cell-free systems⁶⁻¹¹ 2-5A has been shown to activate a ribonuclease. We report here investigations of the sequence specificities of the 2-5A-dependent ribonucleases in extracts of rabbit reticulocytes, mouse ascites tumour cells and human lymphoblastoid cells in conditions of partial digestion using terminally labelled RNA substrates. The enzymes cleaved on the 3'-side of UN sequences to yield UpNp terminated products. Cleavage was observed predominantly at UA and UU sequences.

* Present addresses: Department of Microbiology, Tel Aviv University, Ramat Aviv, Israel (D.H.W.); Department of Pathology, University of Cambridge, UK (J.W.M.); ICRF Laboratories, Lincoln's Inn Fields, London WC2A 3PX, UK (I.M.K.).

Partial nuclease digestions in the presence and absence of tetramer 2-5A were done using rabbit reticulocyte lysates¹² with the 3'-terminally labelled small RNA of the La Crosse virus genome¹³ and influenza virus genome RNAs 7 and 8 as substrates¹⁴. The products were compared by electrophoresis in polyacrylamide gels, with those obtained by specific limited chemical degradations of the RNAs¹⁵. For La Crosse virus sRNA, in the presence of 2-5A, specific cleavages were observed at UN sequences where UA and UU were preferred (Fig. 1a). In this experiment cleavage was obtained at nucleotides 32, 30, 29 and 34, the first of which is at a UA and the others at UU sequences. The precise nature of the products of digestion is not defined by these results and will be described below. Analysis of larger oligonucleotide products of digestion indicated cleavage at residues 54, 58, 69, 100 and 149, in all cases at UA sequences (Fig. 1b, Table 1a). Additional cleavages at nucleotides 59, 66, 73, 123 and 124 were at UU sequences and in the case of nucleotide 72, at a UG. These data are summarized in Table 1. Denaturation of the RNA by boiling before addition to the lysate did not alter the pattern of cleavage (Fig. 1b). In addition, identical results were obtained when 2-5A tetramer, pentamer or oligomers obtained from interferon-treated, virus-infected cells³ were used for activation (data not shown).

The same specificity was observed using influenza virus RNAs 7 and 8 as substrates (Fig. 1c, d; Table 1). With RNA 7 the major sites of degradation were at nucleotides 17, 36 and 64 (Fig. 1c), all at UA sequences. Less extensive cleavage also occurred at nucleotides 23 and 49 (UU sequences), and 53 and 57 (UA sequences). For RNA 8 distinct cleavages were obtained at residues 19, 20, 21, 25 and 26. The major site of cleavage at

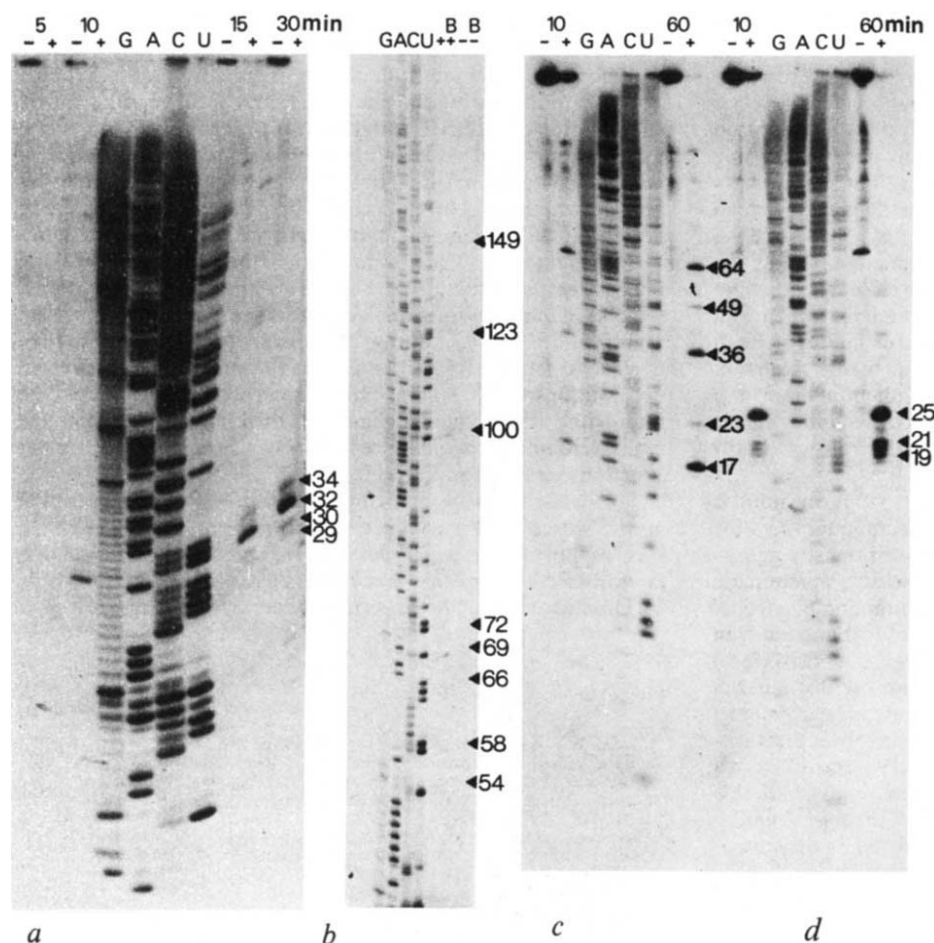
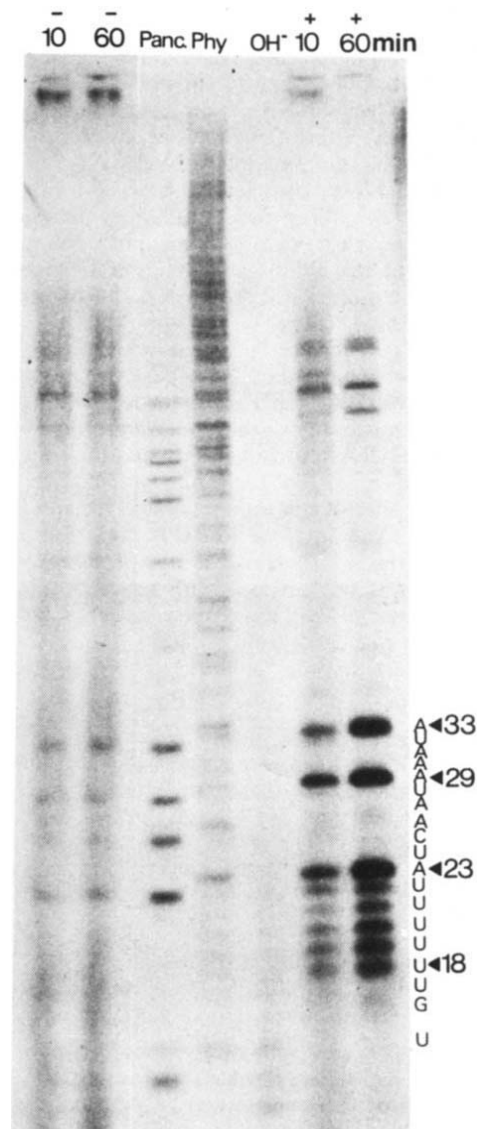


Fig. 1 Comparative electrophoretic analysis of the products of 2-5A-activated RNase activity and those obtained by limited chemical degradation. Virus RNAs were labelled at their 3'-termini with cytidine 5'-[5-³²P]biphosphate and T₄ RNA ligase²¹ and purified by polyacrylamide gel electrophoresis^{13,16}. Assays were performed using rabbit reticulocyte lysates in protein synthetic conditions¹² but in the absence of creatine phosphokinase. Nuclease digestions in the presence (+) or absence (-) of tetramer 2-5A (100 nM) were at 30 °C for up to 60 min. Reactions were terminated by cooling to 0 °C, glycerol was added to 20% w/v and aliquots (10 µl) were applied to polyacrylamide gels containing 7 M urea, 100 mM Tris-borate pH 8.3, and 2.5 mM EDTA. 3'-Labelled RNAs specifically degraded as described by Peattie¹⁵ were applied to adjacent lanes, labelled GACU. *a*, Autoradiogram of a 20% polyacrylamide gel (40 × 40 × 0.05 cm) containing the products of cleavage of La Crosse sRNA after incubation for 5, 10, 15 and 30 min. *b*, Autoradiogram of a 12% polyacrylamide gel (100 × 20 × 0.05 cm) containing the digestion products of La Crosse sRNA after 10 min incubation. In the samples labelled B the RNA substrate was incubated in a boiling water bath for 1 min immediately before addition to the digestion mixtures. *c*, *d*, Autoradiograms of 12% polyacrylamide gels (40 × 40 × 0.05 cm) containing the digestion products of X31 influenza virus RNAs 7 (*c*) and 8 (*d*) after incubation for 10 and 60 min. The numbers indicate the positions of cleavage of the RNA molecules by the 2-5A-dependent nuclease.

grams of 12% polyacrylamide gels (40 × 40 × 0.05 cm) containing the digestion products of X31 influenza virus RNAs 7 (*c*) and 8 (*d*) after incubation for 10 and 60 min. The numbers indicate the positions of cleavage of the RNA molecules by the 2-5A-dependent nuclease.



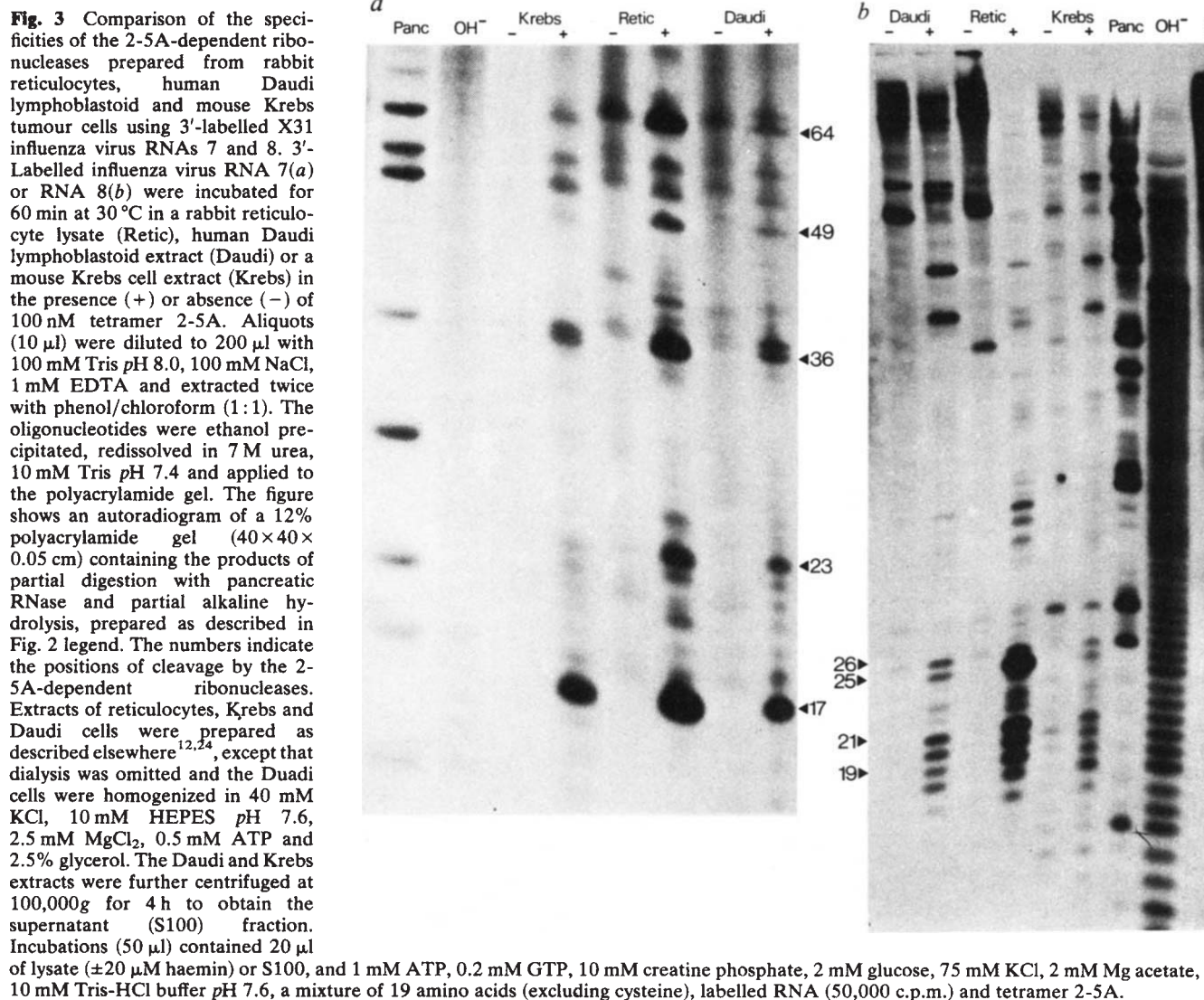
prepared from mouse Krebs ascites tumour cells and human lymphoblastoid (Daudi) cells and their enzyme activities compared with that of a rabbit reticulocyte lysate. Using 3'-labelled influenza virus RNA 7, similar patterns of cleavage were obtained with all three cell extracts: the 2-5A-dependent cleavages seen using the reticulocyte lysate (Fig. 1c) at nucleotides 17, 36 and 64 were also clearly observed using the mouse and human cell extracts (Fig. 3a). Similarly, with influenza virus RNA 8 (Fig. 3b), cleavages at identical positions (nucleotides 19, 20, 21, 25 and 26), although to different extents at nucleotide 25, were obtained in the presence of 2-5A using extracts from cells of all three types.

We conclude from these results that in extracts from cells of several different types, the 2-5A-dependent ribonuclease specifically cleaves at UpN sequences yielding 3'-phosphorylated oligonucleotides which terminate in the sequence UpNp. With the RNA substrates examined the sequences UA and UU were preferred. For the rabbit reticulocyte lysates the most susceptible sites of cleavage were at UA sequences, particularly sequences rich in U residues 32 and 25 in Fig. 1a, d; bold figures in Table 1a, c). However, the preference for sites in U-rich regions was not obvious with the enzymes from the mouse and human cells. This emphasizes that although the 2-5A-dependent nucleases from different sources show similar specificities (Fig. 3), they may not be identical. Furthermore, the

reticulocyte enzyme is known to be unusual in that it requires the tetramer (rather than the trimer) or higher oligomers of 2-5A for full activity³.

Crude extracts were used in these experiments; thus it is possible that more than one nuclease is involved in any given extract. However, studies in progress on the purification of the nuclease and its specific interaction with 2-5A provide results against this.

Cleavage was not obtained at all UA and UU sequences. The reason for this is not known, but it is likely that the specificity is also influenced by the secondary structure of the substrate. In the parts of La Crosse sRNA and influenza RNA 7 examined, the only UA sequences at which cleavage was not observed were in regions of the molecule which may be double stranded^{13,16}. However, the specificity of the nuclease reported does not simply reflect an unusual structure for the viral RNAs because the same specificity for UpN sequences was also observed with ribosomal RNA. Limited cleavages of rRNA have been observed in interferon-treated, SV40 virus-infected cells¹⁷ and in cells into which 2-5A has been deliberately introduced¹⁸. Because specific rRNA cleavages can inactivate ribosome function (for example, those produced by colicin E3 in prokaryotes¹⁹ or by reticulocyte RNase M²⁰) it will be of considerable interest to determine whether rRNA cleavages are a common feature of interferon treatment and virus infection and whether



these or other nucleolytic cleavages in such cells show the same specificity as that reported here.

Since submission of this manuscript P. Lengyel and his colleagues (personal communication) have observed that of the four different ribohomopolymers poly(A), (U), (G) or (C), only poly(U) is cleaved by the 2-5A-dependent nuclease.

D. H. W. was the recipient of an EMBO fellowship. We thank Jack Obijeski for La Crosse virus RNA, Helen Donis-Keller for the *Physarum* RNase and Debora Peattie for communicating the chemical sequencing procedures.

Received 20 August; accepted 27 October 1980.

- Kerr, I. M. & Brown, R. E. *Proc. natn. Acad. Sci. U.S.A.* **75**, 256-260 (1978).
- Stark, G. R., Dower, W. J., Schimke, R. J., Brown, R. E. & Kerr, I. M. *Nature* **278**, 471-473 (1979).
- Williams, B. R. G., Golgher, R. R., Brown, R. E., Gilbert, C. S. & Kerr, I. M. *Nature* **282**, 582-586 (1979).
- Hovanessian, A. G., Wood, J., Meurs, E. & Montagniers, L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 326-326S (1979).
- Williams, B. R. G., Golgher, R. R. & Kerr, I. M. *FEBS Lett.* **105**, 47-52 (1979).
- Clemens, M. J. & Williams, B. R. G. *Cell* **13**, 565-572 (1978).
- Baglioni, C., Minks, M. A. & Maroney, P. A. *Nature* **273**, 684-687 (1978).
- Williams, B. R. G., Kerr, I. M., Gilbert, C. S., White, C. N. & Ball, L. A. *Eur. J. Biochem.* **92**, 455-462 (1978).
- Slattery, E., Ghosh, N., Samanta, H. & Lengyel, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4778-4782 (1979).
- Schmidt, A. *et al.* *FEBS Lett.* **95**, 257-264 (1978).
- Ball, L. A. & White, C. N. *Virology* **93**, 348-356 (1979).
- Williams, B. R. G., Gilbert, C. S. & Kerr, I. M. *Nucleic Acids Res.* **6**, 1335-1350 (1979).
- Obijeski, J. F., McCauley, J. & Skehel, J. J. *Nucleic Acids Res.* **8**, 2431-2438 (1980).
- Scholtissek, C. *Curr. Topics Microbiol. Immun.* **80**, 139-169 (1978).
- Peattie, D. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1760-1764 (1979).
- Skehel, J. J. & Hay, A. J. *Nucleic Acids Res.* **5**, 1207-1219 (1978).
- Revel, M., Kimchi, A., Schmidt, A., Shulman, L. & Chernajovsky, J. *Regulation of Macromolecular Synthesis* (eds Koch, G. & Richter, D) 284-302 (Academic, New York, 1979).
- Hovanessian, A. G. & Wood, J. N. *Virology* **101**, 81-91 (1980).
- Bowman, C. B., Sidikaro, J. & Nomura, M. *Nature new Biol.* **234**, 133-137 (1971).
- Wreschner, D., Melloul, D. & Hertzberg, M. *Eur. J. Biochem.* **89**, 341-352 (1978).
- England, T. E. & Uhlenbeck, O. C. *Biochemistry* **17**, 2069-2076 (1978).
- Donis-Keller, H., Maxam, A. & Gilbert, W. *Nucleic Acids Res.* **4**, 2527-2538 (1977).
- Rommelaere, J., Donis-Keller, H. & Hopkins, N. *Cell* **16**, 43-50 (1979).
- Kerr, I. M., Brown, R. E., Clemens, M. J. & Gilbert, C. S. *Eur. J. Biochem.* **69**, 551-561 (1976).

Removal of O^6 -methylguanine from DNA of normal and xeroderma pigmentosum-derived lymphoblastoid lines

Robert Sklar & Bernard Strauss*

Department of Microbiology, The University of Chicago, 920 East 58th Street, Chicago, Illinois 60637

The ability to excise (repair) UV-induced pyrimidine dimers in *Escherichia coli* is not related to its ability to remove *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG)-induced O^6 -methylguanine (O^6 -MeG) from DNA¹. It was therefore surprising that certain xeroderma pigmentosum cell lines, deficient in dimer excision, were also unable to remove O^6 -MeG¹⁻³. We find that removal of O^6 -MeG occurs rapidly with a half life of less than 1 h. Two cell types can be distinguished: *mex*⁺, which remove O^6 -MeG residues produced by incubation with 0.5 μ g ml⁻¹ MNNG, and *mex*⁻ cells, which are unable to remove the adduct. Xeroderma pigmentosum-derived lymphoblastoid lines of complementation groups A, C or D may be either *mex*⁺ or *mex*⁻. The biochemical mechanism for the removal of O^6 -MeG in human cells is distinct from the excision of adducts produced by compounds such as *N*-acetoxy-*N*-2-acetylaminofluorene (AAAF) or by UV irradiation but it is not clear whether the distinction between *mex*⁺ and *mex*⁻ lines is genetic or epigenetic.

Treatment of cells or DNA with MNNG produces a variety of adducts in the DNA^{4,5}. Of these, O^6 -MeG is particularly important because it seems to lead to increased mutagenesis in

prokaryotes and has been implicated as a potentially carcinogenic lesion⁶⁻⁸. Similar importance has been ascribed to the O^6 -ethylguanine adducts produced by certain ethylating agents.

Cells have evolved systems for dealing with the O^6 -alkylguanine lesion although the detailed biochemistry of the reaction remains to be determined^{9,10}. The removal systems are unequally distributed in mammalian tissues: liver is efficient in removing O^6 -alkylguanine compared with brain tissue which is inefficient¹¹. It has also been observed that simian virus 40 (SV40)-transformed fibroblasts from patients with xeroderma pigmentosum are deficient in the ability to remove O^6 -alkylguanine^{2,3} and this observation has been extended to lymphoblastoid cells derived from an individual with xeroderma pigmentosum¹. To determine whether the inability to remove O^6 -MeG was genetically associated with the xeroderma characteristic, we decided to survey a variety of lymphoblastoid lines, including xeroderma heterozygotes, and determine whether they could remove the O^6 -methyl adduct.

We first determined the kinetics of O^6 -MeG and 3-methyladenine (3-MeA) removal from the DNA of Raji human

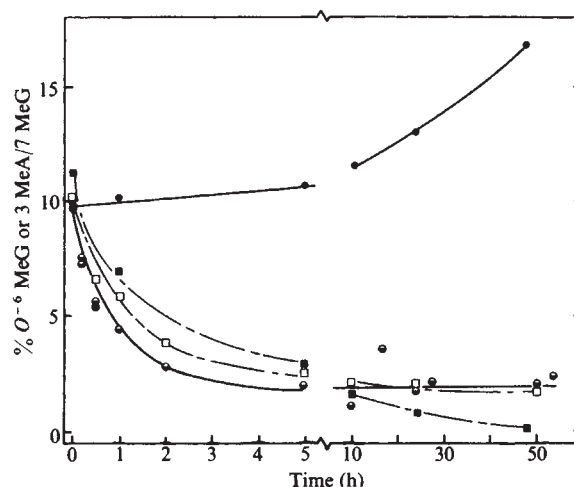


Fig. 1 Kinetics of O^6 -MeG and 3-MeA removal in Raji and H2BT. Two litres of cells ($\sim 8 \times 10^5$ cells ml⁻¹) grown in spinner cultures were centrifuged and resuspended in 90 ml (Raji) or 60 ml (H2BT) of prewarmed SSC (0.15 M NaCl+0.015 M sodium citrate) at 37 °C. Aliquots (10 ml) were treated for 2 min at 37 °C with 0.5 μ g ml⁻¹ ³H-MNNG (112 mCi mmol⁻¹, Radiochemicals, Amersham) and 1 ml *N*-acetyl-L-cysteine (100 mM, pH 7.2) (ref. 12). Immediately after treatment a 0-time aliquot was lysed with 1 ml 2% SDS in SSC and frozen; later samples were centrifuged and resuspended in 100 ml RPMI 1640 plus 10% fetal calf serum and 4 mM L-glutamine in a 75 cm² Falcon flask and incubated with mild agitation for the time indicated. Cells were then centrifuged, resuspended in SSC and lysed with SDS as above. Cell lysates (10 ml) were treated for 1 h with 0.5 ml of 20 mg ml⁻¹ heat-treated RNase (Worthington Biochemical) and for 4 h with 1.5 ml of 20 mg ml⁻¹ heat-treated pronase (Calbiochem-Behring). Digested lysates were extracted three times with redistilled phenol saturated with SSC and neutralized with 0.2 volumes of 1 M Tris base. DNA was precipitated by adding 0.1 volume of 2.5 M sodium acetate and 2 volumes of 2-ethoxyethanol, washed twice with ice-cold 100% ethanol and dried. The DNA was dissolved in 100 μ l of 0.05 M ammonium formate with 0.02% sodium azide, pH 7.5, and hydrolysed at 70 °C for 20 min after adding 0.2 volumes of 1 M HCl and a 50 μ l mixture of 3-MeA (2 mg ml⁻¹, Cyclo Chemical), 7-MeG (5 mg ml⁻¹, Sigma), and O^6 -MeG (5 mg ml⁻¹) (O^6 -MeG, obtained from Dr D. Ludlum, was hydrolysed to yield the free base). The hydrolysate was spotted over 6 cm of a 0.5-mm cellulose MN300 thin-layer chromatography plate pre-chromatographed with solvent (Analtch) and developed in 2-propanol:28% NH₄OH:H₂O (7:2:1) (ref. 11). UV absorbing bands corresponding to the above markers and guanine were removed from the plate with a razor blade and placed in scintillation vials, then 2 ml of 0.05 M ammonium formate, 0.02% sodium azide with 0.1 volume of 1 M HCl were added. Aquasol was added to the radioactive bands and the vials were vortexed and counted after several hours of equilibration. The guanine band was centrifuged twice in a microcentrifuge to remove the cellulose and diluted for absorbancy reading at 260 nm. The amount of DNA (in μ g) was determined by multiplying the total guanine absorbancy by 164, a number determined from the absorbancy of a known amount of guanine and the molar ratio of guanine in DNA. Squares represent ratio of 3-MeA to 7-MeG $\times 100$; circles, ratio of O^6 -MeG to 7-MeG $\times 100$; completely filled symbols, line H2BT; open squares and half-filled circles, lymphoma line Raji.

* To whom correspondence should be addressed.

lymphoma cells because this line is known to remove methylation products from DNA¹. Exponentially growing cells were treated with a pulse of MNNG and at intervals portions of the culture were collected and the alkylation products analysed. Figure 1 shows that *O*⁶-MeG was rapidly removed from the DNA with a half life of less than 1 h at 37 °C, and that at the concentration of MNNG used (0.5 µg ml⁻¹) the ratios of *O*⁶-MeG and 3-MeA to 7-methylguanine (7-MeG) declined at about the same rate. Experiments at higher MNNG concentrations indicate that the capacity for *O*⁶-MeG removal is much more limited than for 3-MeA. The results in Fig. 1 are expressed as a ratio of radioactivity in *O*⁶-MeG to 7-MeG and because

7-MeG is gradually lost from the DNA¹, the inability to remove the *O*⁶-MeG adduct from the DNA results in an increase in this ratio with time of incubation in strain H2BT.

The result that the H2BT line, derived from a spherocytosis patient and heterozygous for thymidine kinase¹³, was deficient in *O*⁶-MeG removal was unexpected, as was the discovery that other non-xeroderma lines were also deficient (Table 1). In a series of control experiments in the Raji line at different MNNG doses we found that the 0-time value for the *O*⁶-MeG/7-MeG ratio was 10.4% and the 0-time 3-MeA/7-MeG ratio was 10.1%. All the lines which remove *O*⁶-MeG have *O*⁶-MeG/7-MeG ratios which are consistent with the data obtained with

Table 1 Removal of *O*⁶-MeG by human lymphoblastoid lines

Strain		Incubation time (h)	7-MeG (d.p.m. per μg DNA)	O ⁶ -MeG/7-MeG × 100	3-MeA/7-MeG × 100	Classification	
Designation	Origin						
a Burkitt's lymphoma lines							
Raji	Burkitt's lymphoma	0	13.52	10.9	9.76	mex ⁺	
		27.5	3.59	1.2	2.15		
Daudi	Burkitt's lymphoma	0.5	2.04	8.0	3.8	mex ⁺	
		10.0	1.99	3.3	2.8		
HR-1	Burkitt's lymphoma	0	—	7.1	10.2	mex ⁺	
		3.0	4.46	0.73	2.5		
b Human lymphoblastoid lines							
H2BT	Thymidine kinase heterozygote derived from a spherocytosis patient	0	10.85	9.7	11.3	mex ⁻	
		24.0	2.93	13.0	1.2		
L 33-6-1	Infectious mononucleosis	0	7.56	10.2	9.47	mex ⁻	
		24.3	2.03	15.4	0.4		
GM 892	Normal female (E. Conod, 1970) from HGMCR*	0	8.34	10.0	9.14	mex ⁻	
		24.0	4.43	12.9	1.4		
GM 621	Normal male established 6/4/74 from HGMCR	0	9.21	9.83	9.85	mex ⁻	
		24.0	4.42	11.7	1.5		
GM 1953	Normal female established 3/17/75 from HGMCR	0	—	8.09	6.62	mex ⁺	
		3.0	—	3.05	3.25		
D 8-5-1	Progeria patient†	0.5	9.05	3.94	7.86	mex ⁺	
		29.0	4.18	0.82	1.54		
D 16-1-1	Werner's Syndrome†	0	11.63	9.20	11.4	mex ⁺	
		25.5	8.43	2.27	1.31		
c Xeroderma pigmentosum-related lines							
		Complementation group					
GM 2250	HGMCR‡	A	0	3.1	8.3	11.7	mex ⁺
			24.0	1.7	1.5	2.2	
GM 2500	HGMCR‡	A	0	11.5	10.0	—	mex ⁺
			24.0	11.6	2.5	—	
GM 2246	HGMCR‡	C	0	12.7	9.4	11.5	mex ⁺
			24.0	7.6	1.6	1.3	
GM 2248	HGMCR‡	C	0	—	10.5	11.3	mex ⁻
			24.0	—	11.8	1.4	
GM 2498	HGMCR‡	C	0	8.1	10.4	11.6	mex ⁻
			24.0	5.1	14.0	1.6	
GM 2253	HGMCR‡	D	0.5	—	3.81	5.82	mex ⁺
			3.0	—	2.02	2.72	
GM 2473	HGMCR‡	D heterozygote mother of 2253	0.5	—	9.78	6.79	mex ⁻
			3.0	—	9.86	3.88	
D 70-1-2	Xeroderma heterozygote (father, family JAGR)†	Unknown	0	8.74	6.85	9.68	mex ⁺
			3.0	9.45	1.6	2.7	
D 70-2-2	Xeroderma heterozygote (mother, family JAGR)†	Unknown	0	—	9.9	10.5	mex ⁺
			23.0	—	3.2	1.7	
D 70-3-1	Xeroderma proband, family JAGR†	Unknown	0	4.25	8.44	8.44	mex ⁺
			3.0	4.67	1.7	3.8	
D 70-3-2	Xeroderma proband†. Independent lymphoblastoid culture.	Unknown	0	5.15	11.9	14.8	mex ⁺
			32.0	2.92	2.5	1.8	
A 64-1-2	Xeroderma heterozygote (mother), MALO family.†	Unknown	0.5	14.4	3.41	8.14	mex ⁺
			29.0	6.06	0.38	1.0	
A 64-2-3	Xeroderma proband, male, MALO family.†	Unknown	0	8.25	7.74	9.27	mex ⁺
			24.0	4.56	0.93	1.2	
D 75-1-2	Xeroderma, male, independent culture from A 64-2-3.†	Unknown	0	6.19	8.79	10.7	mex ⁺
			25.0	4.0	3.0	1.6	
D 75-2-1	Xeroderma patient, MALO family, sister of A 64-2-3.†	Unknown	0	11.75	8.68	9.71	mex ⁺
			25.0	8.67	3.0	1.3	
B 5-3-1	Xeroderma heterozygote, MALO family father.†	Unknown	0	11.48	9.25	10.3	mex ⁺
			25.0	4.94	3.23	1.5	

500 ml of cells ($\sim 8 \times 10^5$ ml⁻¹) grown in spinner culture were collected by centrifugation and resuspended in 20 ml SSC. Aliquots (10 ml) were treated with 0.5 µg ml⁻¹ ³H-MNNG (112 mCi mmol⁻¹), resuspended in RPMI 1640 as described in Fig. 1. legend and incubated for the indicated time. The cells were then collected and lysed and the DNA purified as described in Fig. 1 legend. The Burkitt's lymphoma lines were obtained from P. Gerber (Raji) and G. Klein (Daudi and HR-1). H2BT supplied by W. Thilly; L 33-6-1 from J. Littlefield.

* HGMCR, Human Genetic Mutant Cell Repository, Camden, New Jersey.

† From Earl Henderson.

‡ From A. Andrews.

Table 2 Excision repair in mex⁺ and in a mex⁻ lymphoblastoid line

Strain	Origin	Control	Repair synthesis (c.p.m. μg^{-1}) induced by:											
			AAAF ($\mu\text{g ml}^{-1}$)		AAAF/ control (25 μg)	MMS ($\mu\text{g ml}^{-1}$)			MMS/ control (440 μg)	MNNG ($\mu\text{g ml}^{-1}$)			MNNG/ control (5 μg)	
			12.5	25		110	220			1	2.5	5		
D 70-1-2	JAGR father (mex ⁺)	13.2		128	(9.7)									
D 70-2-2	JAGR mother (mex ⁺)	11.1		138	(12.4)									
D 70-3-1	JAGR proband (mex ⁺)	16.6		22	(1.3)									
D 70-3-2	JAGR proband (mex ⁺) (2nd culture)	12.9		22	(1.7)									
D 70-4-1	JAGR brother	28.2		170	(6.0)									
D 70-5-2	JAGR sister	18.3		136	(7.4)									
D 70-1-2	JAGR father (mex ⁺)	79.0	476	733	(9.3)	192	294	285	(3.6)	129	272	424		(5.4)
GM 621	Normal male (mex ⁻)	37.0	286	398	(10.8)	162	220	179	(4.8)	88	169	243		(6.6)

DNA excision repair was measured by the BND cellulose method¹⁶. Rapidly growing cells were diluted to $4 \times 10^5 \text{ ml}^{-1}$, grown overnight, concentrated to about 10^6 ml^{-1} , and 2 ml of suspension used for determination. Repair synthesis was measured in the presence of 2 mM hydroxyurea and 10^{-6} M fluorodeoxyuridine along with $5 \mu\text{Ci ml}^{-1}$ ^3H -thymidine (15 Ci mmol^{-1}). Results were not corrected for pool size and the ratio of drug-induced counts to the background is given as an indication of the relative repair capacity. Determinations of repair induced by MMS, AAF and MNNG were done separately from the tests of the JAGR family, which were carried out by Mr Dino Vallera.

Raji (Fig. 1). Of 26 lymphoma and lymphoblastoid lines classified for their ability to remove O^6 -MeG, we found 19 removers and 7 non-removers. We call the non-removers mex^- (methyl excision minus) for convenience but the terminology has no genetic significance.

Mex^- strains include three Burkitt's lymphoma lines, lines derived from xeroderma patients of complementation groups A, C and D, and representatives of tx xeroderma families of unknown complementation group. All xeroderma lines were tested in this laboratory for repair synthesis induced by AAFF and found to be repair negative compared with controls including the xeroderma heterozygotes. As the lymphoblastoid cultures are continuous lines, this result indicates that transformation does not affect repair deficiency. The seven mex^- lines we have identified include a line derived from an infectious mononucleosis patient, a spherocytosis patient, a xeroderma heterozygote and a normal male and female culture.

The dose of MNNG used in these experiments ($0.5 \mu\text{g ml}^{-1}$) reduced the dose of survival of mex^+ Raji cells to $\sim 20\%$, as determined in a cloning experiment. Inability to remove O^6 -MeG adducts after treatment at $0.5 \mu\text{g ml}^{-1}$ may indicate sensitivity of MNNG at lower concentrations for mex^- strains. Most lymphoblastoid lines cannot be efficiently cloned in soft agar but their growth can be followed by cell counts in liquid medium¹⁴. We have observed (Fig. 2) that mex^- cells are more sensitive to growth inhibition at low MNNG doses than are removal-competent strains (mex^+). The greater sensitivity of mex^- strains was not found when treated with the carcinogen AAAF whose adducts are removed by UV-like (nucleotide) excision¹⁵. The mex^- strain GM 892 was also more sensitive to growth inhibition by 3 or $5 \mu\text{g ml}^{-1}$ *N*-ethyl-*N'*-nitro-*N*-nitrosoguanidine (ENNG) than the mex^+ D70-1-2, and ENNG pretreatment ($2.5 \mu\text{g ml}^{-1}$) 2 h before MNNG challenge eliminated the capacity of the mex^+ Raji line to remove O^6 -MeG, suggesting that the mex system is involved in the metabolism of ethyl adducts.

Both mex^+ and mex^- strains respond to treatment with methyl methanesulphonate (MMS), MNNG or AAF with increased DNA excision-repair synthesis (Table 2) and estimation of the relative activities by comparison of the ratio of repair synthesis with the controls (Table 2) indicates that the mex^+ D70-1-2 and mex^- GM 621 give an almost equivalent repair synthesis response. Inability to respond to AAF by increased repair synthesis is not related to the ability to remove O^6 -MeG in lymphoblastoid cells (Tables 1 and 2).

The result that mex⁻ strains respond to treatment with MNNG by increased repair synthesis seems to indicate that MNNG produces a variety of lesions, some of which (for example, 3-MeA) are removed, producing apurinic sites which can be repaired by an intact AP (apurinic/aprimidinic) repair system, and either that *O*⁶-MeG removal is a minor reaction which does not show up in the repair measurements at the high MNNG dose used, or that removal of the *O*⁶-methyl group is

not accompanied by base removal as suggested by *in vitro* studies^{9,10}.

We do not know the genetic basis for the mex^- characteristic. There could be a major mex polymorphism in human populations and this hypothesis could account for a mex^+ xeroderma line and a mex^- heterozygous line from the same family (Table 1). Alternatively, some epigenetic process might occur in cultured cells resulting in a mex^+ to mex^- conversion. Such processes are known to occur in normal development because different organs of the same animal (for example, brain and liver) can have a different mex characteristic¹¹. If the process is epigenetic, it is possible that functions other than the ability to remove O^6 -MeG would be simultaneously affected.

Our mex^+ and mex^- strains have properties like those studied by Day *et al.*^{17,18} who reported strains which differ in their ability to reactivate MNNG-treated adenovirus. We have found that a pair of Day's mer^- and mer^+ are mex^- and mex^+ respectively and this observation has been confirmed (Day, personal communication). Separation of the mex characteristic from xeroderma phenotype also agrees with the finding that only a SV40-transformed xeroderma cell line was hypersensitive to sister chromatid induction by the alkylating agent, ethyl methanesulphonate¹⁹.

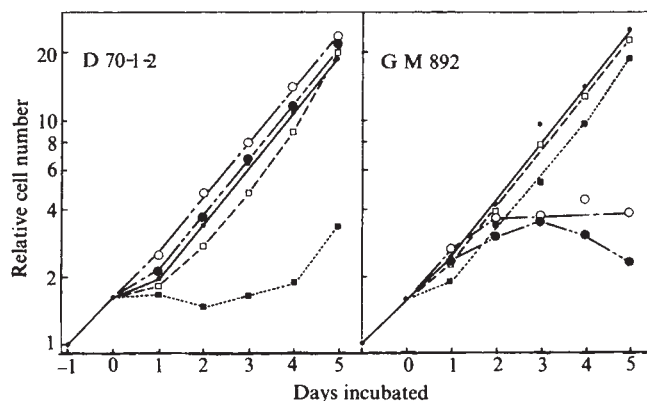


Fig. 2 Effect of MNNG and AAAF on the growth of the mex⁺ (D70-1-2) and mex⁻ (GM 892) lymphoblastoid lines. Cultures were maintained in a rapidly growing state by daily dilution to $4 \times 10^5 \text{ ml}^{-1}$ in RPMI 1640 + 16% fetal calf serum. At day '0', the cells were diluted to $4 \times 10^5 \text{ ml}^{-1}$ and treated with AAAF or MNNG for 1 h at 37°C. The cells were then collected by centrifugation, resuspended in medium at $4 \times 10^5 \text{ ml}^{-1}$ and incubated at 37°C. Cultures were counted daily using a Coulter counter and the population density was adjusted to $4 \times 10^5 \text{ ml}^{-1}$. The daily increase in population size was used to calculate a relative cell number assuming no dilution. AAAF was dissolved in dimethyl sulphoxide, which was added in equal amounts to all cultures. Small ●, control; large ○, MNNG (0.1 $\mu\text{g ml}^{-1}$); large ●, MNNG (0.2 $\mu\text{g ml}^{-1}$); □, AAAF (4 $\mu\text{g ml}^{-1}$); ■, AAAF (8 $\mu\text{g ml}^{-1}$).

We thank Dr Earl Henderson for supplying mex⁺ strains. This work was supported by NCI contract NO1-CP-85669, and by grants from the NCI (CA 14599) and the Department of Energy (EY75 S-02 2040). R.S. was a trainee on a NCI-sponsored training grant in environmental biology (CA 09273).

Received 7 August; accepted 7 November 1980.

- Altamirano-Dimas, M., Sklar, R. & Strauss, B. *Mutat. Res.* **60**, 197–206 (1979).
- Goth-Goldstein, R. *Nature* **267**, 81–82 (1977).
- Bodell, W., Singer, B. & Cleaver, J. *Nucleic Acids Res.* **6**, 2819–2829 (1979).
- Warren, W. & Lawley, P. *Carcinogenesis* **1**, 67–78 (1980).
- Singer, B. *Nature* **264**, 333–339 (1976).
- Loveless, A. *Nature* **223**, 206–207 (1969).
- Schendel, P. & Robins, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 6017–6020 (1978).
- Sklar, R. & Strauss, B. *J. molec. Biol.* **143**, 343–362 (1980).
- Karran, P., Lindahl, T. & Griffin, G. *Nature* **280**, 76–77 (1979).
- Pegg, A. *Biochem. biophys. Res. Commun.* **84**, 166–173 (1978).
- Goth, R. & Rajewsky, M. *Proc. natn. Acad. Sci. U.S.A.* **71**, 639–643 (1974).
- Lawley, P. & Thatcher, C. *Biochem. J.* **116**, 693–707 (1970).
- Skopec, T., Liber, H., Penman, B. & Thilly, W. *Biochem. biophys. Res. Commun.* **84**, 411–416 (1978).
- Higgins, N. P. & Strauss, B. *Cancer Res.* **39**, 312–320 (1979).
- Regan, J. & Setlow, R. *Cancer Res.* **34**, 3318–3325 (1974).
- Scudiero, D., Henderson, E., Norin, A. & Strauss, B. *Mutat. Res.* **29**, 473–488 (1975).
- Day, R. & Ziolkowski, C. *Nature* **279**, 797–799 (1979).
- Day, R., Ziolkowski, C., Scudiero, D., Meyer, S. & Mottern, M. *Carcinogenesis* **1**, 21–32 (1980).
- Heddle, J. & Arlett, C. *Mutat. Res.* **72**, 119–125 (1980).

Assembly units of clathrin coats

Ernst Ungewickell & Daniel Branton

The Biological Laboratories, 16 Divinity Avenue,
Harvard University, Cambridge, Massachusetts 02138

Clathrin, a polypeptide of molecular weight (MW) 180,000, is the main constituent of the polygonal network that forms the coat of coated pits and vesicles¹; these vesicles play a part in intracellular transport between membranous organelles^{2,3}. This function involves specific recognition of target membranes as well as fusion and fission events that must be coordinated with the assembly, partial disassembly or reorganization of the clathrin coats. To understand these interactions on a molecular level, information about the structure of clathrin and the interactions of clathrin with itself and other proteins is required. Here we show that purified clathrin coats dissociate reversibly into triskelions, structures composed of three usually bent, rather flexible legs radiating from a centre. We have determined the molecular weight of these triskelions and conclude that they contain trimers of clathrin together with about three light molecular weight polypeptide chains.

Coated vesicles were isolated from fresh calf brains according to the procedure of Keen *et al.*⁴ and the clathrin coats were dissociated in 2 M urea⁵. Two successive gel filtrations separated clathrin from polypeptides with apparent molecular weights of ~90,000–125,000 and 50,000–60,000, but failed to remove two light chains with apparent molecular weights of ~36,000 and 33,000 (Fig. 1). Experiments in which this column-purified clathrin was reacted with the cleavable cross-linking reagent 3,3'-dithiobis-propionimidate confirmed Pearse's⁶ suggestion that these two light chains are bound to clathrin (results not shown). From the relative intensities of the Coomassie Brilliant Blue staining, we estimated that there is approximately one light chain per clathrin polypeptide. This ratio did not vary significantly between preparations although we did observe that the clathrin light chains are very sensitive to proteolysis.

It has been reported that antibodies to two polypeptides in the 30,000–40,000 MW range (probably the light chains) cross-react with brain tropomyosin⁷. We compared the electrophoretic mobilities of brain tropomyosin and clathrin light chains in SDS-polyacrylamide gels and in SDS-polyacrylamide gels containing urea. Tropomyosins from skeletal muscle and calf brain show anomalously low electrophoretic mobility in urea/SDS-polyacrylamide gels⁸. In these conditions brain tropomyosin with an apparent molecular weight of 30,000 in SDS

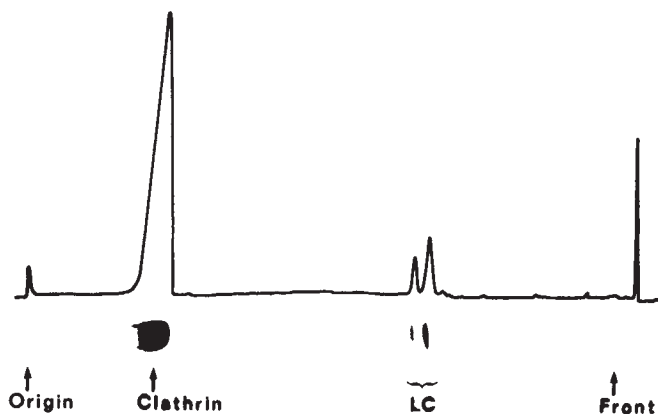


Fig. 1 Polyacrylamide gel electrophoresis of purified dissociate clathrin coats. Coated vesicles were prepared according to the procedure of Keen *et al.*⁴ except that all solutions contained 0.005% phenylmethylsulphonyl fluoride (PMSF). To dissociate the coats, purified coated vesicles were dialysed overnight at 4°C against 500 volumes of 30 mM Tris-HCl, 20 mM NaCl, 1 mM EDTA, 0.2 mM dithiothreitol (DTT), 2 M urea, 0.02% NaN₃ and 0.005% PMSF, pH 7.5. Vesicles were removed by centrifugation at 100,000g for 1 h. The supernatant, which contained large amounts of protein in addition to clathrin, was chromatographed at 4°C on a Sepharose 4B-Cl gel filtration column (90×2.5 cm), which was equilibrated with dialysis buffer. Clathrin eluted together with two low-MW polypeptides (light chains) and small amounts of polypeptides with apparent molecular weights of ~100,000 and 55,000 (molecular weights were estimated from the relative mobilities of polypeptides in SDS-polyacrylamide gels) which were completely removed by a second gel filtration in triethanolamine. The clathrin-containing fractions from the first column were concentrated by precipitation with 50% ammonium sulphate and immediately dissolved in 50 mM triethanolamine, 20 mM NaCl, 1 mM EDTA, 0.2 mM DTT and 0.02% NaN₃, pH 8.3. After dialysis against the same buffer, the protein was chromatographed in the same medium at 4°C on a Sephadex 4B-Cl gel filtration column (100×1.5 cm). The polypeptide composition of the final pooled clathrin-containing fractions was analysed in SDS-polyacrylamide gradient gels (6–15%) using the discontinuous buffer system of Laemmli⁹. The gel was scanned with a Joyce Loeb microdensitometer. LC, clathrin light chains.

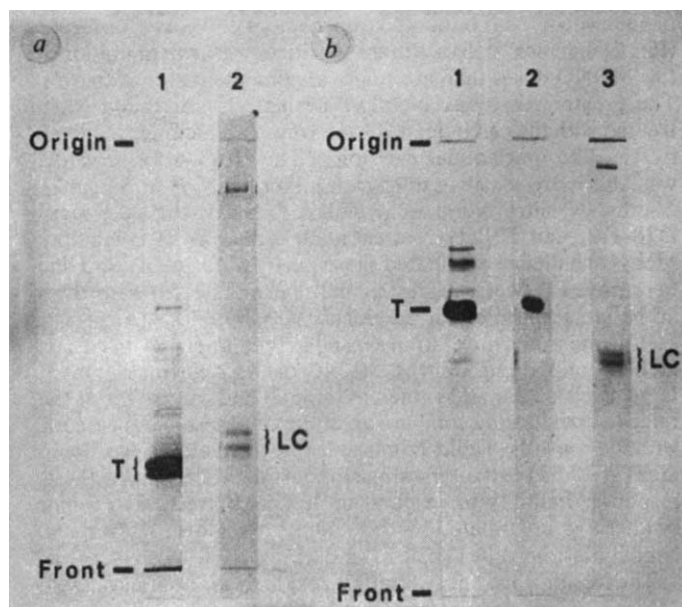


Fig. 2 Comparison between the electrophoretic behaviour of brain tropomyosin and clathrin light chains in SDS- and urea/SDS-polyacrylamide gels. *a*, 10% SDS-polyacrylamide gel⁸: lane 1, brain tropomyosin (T); lane 2, brain clathrin light chains (LC). *b*, urea/SDS-polyacrylamide gel: lane 1, brain tropomyosin; lane 2, skeletal muscle actin; lane 3, clathrin light chains.

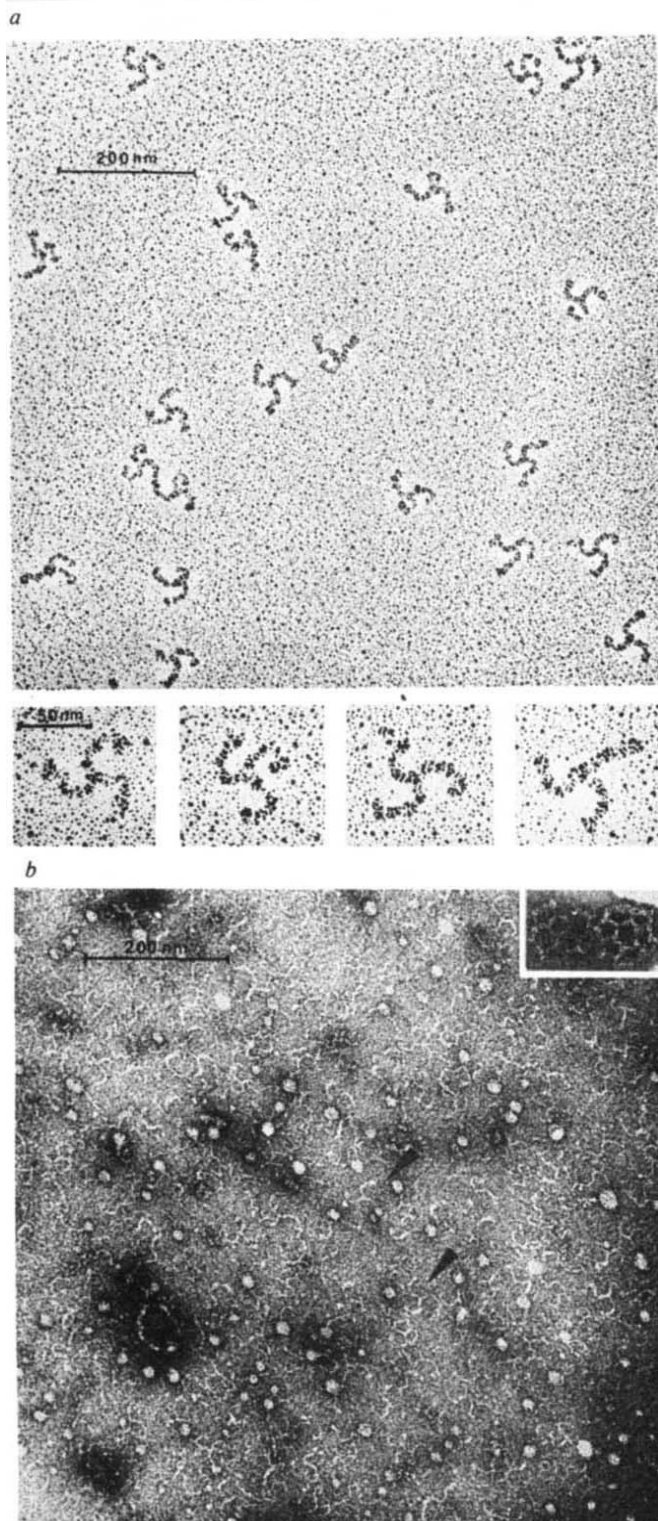


Fig. 3 Electron micrographs of 8.4 s clathrin triskelions, after *a*, rotary shadowing and *b*, negative staining with uranyl acetate. For rotary shadowing, dissociated clathrin coats ($20 \mu\text{g ml}^{-1}$) in 50 mM triethanolamine, pH 8.3, were mixed at room temperature with two volumes of glycerol and sprayed onto freshly cleaved mica. The molecules were then dried at room temperature under vacuum in a Balzers freeze-etch device, rotary-shadowed at an angle of 4.5° with Pt/C and coated from above with pure carbon (for details of method see ref. 11). For negative staining dissociated clathrin coats in 0.3 mM Tris-HCl, pH 8.0 ($20 \mu\text{g ml}^{-1}$) were applied to Formvar- and carbon-coated grids and negatively stained with 1% uranyl acetate. Arrows in *b* mark structures which resemble those seen in *a*. Inset in *b* shows coats reassembled from purified triskelions after dialysis against buffer A. All specimens were examined in a Philips EM 301 electron microscope at 80 kV. *a*, $\times 90,000$ and $180,000$; *b*, $\times 90,000$.

gels has almost the same electrophoretic mobility as actin (43,000 MW) (ref. 8 and Fig. 2). In contrast, the clathrin light chains which ran with apparent molecular weights of 36,000 and 33,000 in SDS did not show a low electrophoretic mobility in SDS gels containing urea (Fig. 2). From this observation and from results of one-dimensional peptide maps (not shown) we conclude that the light chains and brain tropomyosin are not closely related proteins.

Analytical ultracentrifugation showed that the purified clathrin-light chain complex had a sedimentation coefficient, $s_{20,w}$, of 8.4 in 50 mM triethanolamine, 20 mM sodium chloride, 1 mM EDTA at pH 8.3. On dialysis against 0.1 M 2-(*N*-morpholino)-ethanesulphonic acid, 0.5 mM magnesium chloride, 0.2 mM EGTA and 0.002% sodium azide, pH 6.5 (buffer A), the 8.4 s material formed typical coats or baskets with relatively uniform diameters of $\sim 800 \text{ \AA}$ (Fig. 3*b* inset). In several reassembly experiments, 75–93% of the total protein was recovered in the pelleted baskets after centrifugation at $100,000g$ for 1 h. The polypeptide composition of the pellets was identical to that of the starting material.

After rotary shadowing or negative staining, the 8.4 s material appeared as triskelions with a mean leg length of $445 \text{ \AA} \pm 23$ (s.d.). The legs in most of the molecules showed a preferred direction of curvature. When the curvature was sharp, it was located $\sim 190 \text{ \AA}$ from the centre point where the three legs join. At higher magnification each leg seems to be composed of two strands which are twisted around each other and which seem to form a hairpin loop at the end.

The molecular weight of triskelions was determined by electron microscopy¹² and sedimentation equilibrium. For electron microscopy we mixed triskelions at a known weight concentration (determined by colorimetric micro-Kjeldahl analysis¹³ using a nitrogen factor of 6.33 (ref. 14)) with myosin (at a known weight concentration), and counted the ratio of triskelions to myosin molecules in electron micrographs of droplets that had been dried and rotary shadowed. The estimated molecular weight was $627,000 \pm 55,000$ (s.d.; total number of particles counted was $\sim 2,000$). For sedimentation equilibrium the high-speed Yphantis method was used¹⁵. An apparent molecular weight of 630,000 was calculated from the plot in Fig. 4.

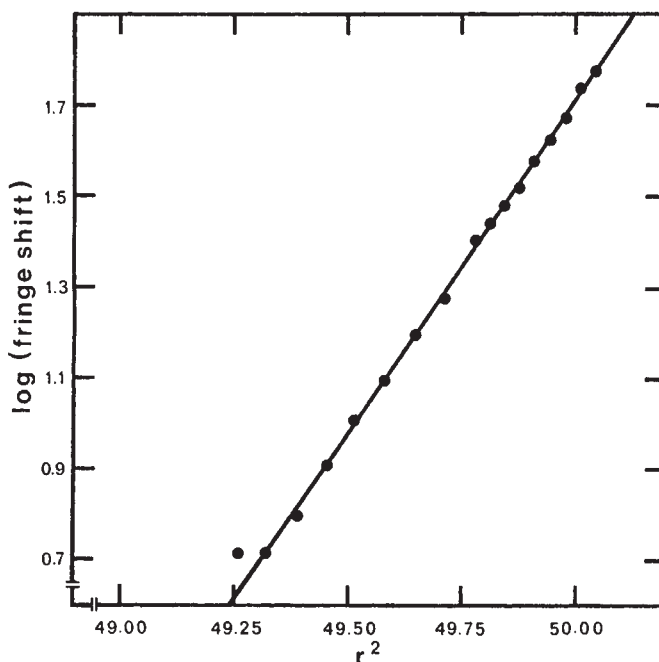


Fig. 4 Sedimentation equilibrium distribution of triskelions. Sedimentation equilibria were performed by the high speed method¹⁵, with the use of the Rayleigh interference optical system in a Spinco model E ultracentrifuge at 9,341 r.p.m. and 9.1°C . The protein concentration was 0.5 mg ml^{-1} , and the partial specific volume calculated from the amino acid composition was $0.75 \text{ cm}^3 \text{ g}^{-1}$ (ref. 14).

Assuming a 180,000 MW for a single clathrin polypeptide¹, the most straightforward interpretation of our data is that triskelions comprise three clathrin monomers and probably three light chains. We do not know whether a single triskelion contains both the 36,000 and 33,000 MW light chains. Because triskelions are stable molecular structures observed in various conditions, it is probable that other investigators studying the properties of dissociated clathrin coats have been dealing with triskelions (for example, the 8.8 s preparation of Puszkun and coworkers^{16,17}).

According to Crowther *et al.*¹⁸, a clathrin coat which is constructed of 8 hexagons and 12 pentagons (36 vertices) contains approximately 100 clathrin monomers. This suggests that each vertex in the coat is the centre of a triskelion. A triskelion leg is long enough to span two edges of the polygons

that make up the coat^{10,18} but other arrangements in which a portion of each leg could extend out of the coat towards the contained vesicle cannot be excluded.

We thank F. Marty, A. Verkleij, D. Caspar, S. Lehrer and D. Siegel for encouragement and discussions, W. Gratzer for help with the hydrodynamic measurements (which were performed at the MRC Cell Biophysics Unit in London), and Ken Giedd for technical assistance. E.U. was supported by a Training Grant from the Deutsche Forschungsgemeinschaft; the research was supported by NIH grant GM 20396.

Note added in proof: T. Kirchhausen and S. C. Harrison (Harvard) have recently shown by cross-linking experiments that the 8 s species contains three heavy chains, each in close contact with a light chain (manuscript submitted for publication).

Received 23 September; accepted 18 November 1980.

1. Pearse, B. M. F. *J. molec. Biol.* **97**, 93–98 (1975).
2. Goldstein, J. L., Anderson, R. G. W. & Brown, M. S. *Nature* **279**, 679–685 (1979).
3. Pearse, B. *Trends biochem. Sci.* **5**, 131–134 (1980).
4. Keen, J. H., Willingham, M. C. & Pastan, J. H. *Cell* **16**, 303–312 (1979).
5. Schook, W., Ores, C. & Puszkun, S. *J. Cell Biol.* **15**, 119a (1977).
6. Pearse, B. M. F. *J. molec. Biol.* **126**, 803–812 (1978).
7. Puszkun, S., Maimon, J. & Schook, W. *J. Cell Biol.* **83**, 293a (1979).
8. Bretscher, A. & Weber, U. *FEBS Lett.* **85**, 145–148 (1978).
9. Laemmli, U. U. *Nature* **227**, 680–685 (1970).
10. Heuser, J. *J. Cell Biol.* **84**, 560–583 (1980).

11. Tyler, J. M. & Branton, D. *J. ultrastruct. Res.* **71**, 95–102 (1980).
12. Williams, R. C. & Backus, J. *J. Am. chem. Soc.* **71**, 4052–4057 (1949).
13. Jaenicke, L. *Analyt. Biochem.* **61**, 623–627 (1974).
14. Pearse, B. M. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1255–1259 (1976).
15. Yphantis, D. A. *Biochemistry* **3**, 297–317 (1964).
16. Schook, W., Puszkun, S., Bloom, W., Ores, C. & Kochwa, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 116–120 (1979).
17. Bloom, W. S., Schook, W., Feagson, E., Ores, C. & Puskin, S. *Biochim. biophys. Acta* **598**, 447–455 (1980).
18. Crowther, R. A., Finch, J. T. & Pearse, B. M. F. *J. molec. Biol.* **103**, 785–798 (1976).

Disulphide bonds of haemagglutinin of Asian influenza virus

Michael Waterfield & Geoffrey Scrace

Imperial Cancer Research Fund Laboratories, Lincoln's Inn Fields, London WC2A 3PX, UK

John Skehel

National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The complete amino acid sequences of the haemagglutinins (HAs) of one strain of each of the Hav1 and H2 subtypes and of three strains of the H3 subtype of influenza virus have been established^{1–3}. A large external fragment (BHA) of HA can be released from virus particles using the protease bromelain⁴ that cleaves the protein close to its carboxyl terminus, which is inserted in the virus membrane (see Fig 1). Here we show that a single disulphide bond joins the two component polypeptide chains BHA₁ and BHA₂ and establish the location of five intrachain disulphides. These disulphide bonds have been located in the three-dimensional structure of HA which is known to 3 Å resolution^{5,6}.

Sequence analysis of HA of infectious A/Japan/305/57(H2N1) shows that the subunit contains 546 amino acids, 16 of which are half-cystine residues³. Analyses of BHA₁ and BHA₂ calculated as 324 and 170 residues respectively⁷ reveal 9.1 and 2.9 half-cystine residues in these polypeptides. The precise location of the bromelain cleavage site(s) is unknown because the enzyme does not cleave the HA of this strain at a single site. However, cleavage is known to be C-terminal to half-cystine 473 in BHA₂ (ref. 8). Established sequence data suggest that four of the 16 half-cystines are in the C terminus of HA₂, which is removed by bromelain treatment. Alkylation of half-cystine residues of BHA in 6 M guanidine hydrochloride using ¹⁴C-iodoacetamide failed to reveal any free sulphhydryl groups, thus all the 12 half-cystines in BHA must exist in disulphide bonds.

To locate the half-cystines linked by disulphide bonds, fragments containing cystine residues were purified. BHA was first cleaved with cyanogen bromide and fragments fractionated by size using gel permeation HPLC (Fig. 1). Amino acid analyses of performic acid-oxidized aliquots showed that peaks A, B, C and D contained cysteic acid. SDS polyacrylamide gel analysis showed that peak A contained a mixture of partial cleavage products. Aliquots of material in the incompletely resolved peaks B and C were fully reduced with dithiothreitol and component polypeptides analysed by gel permeation HPLC

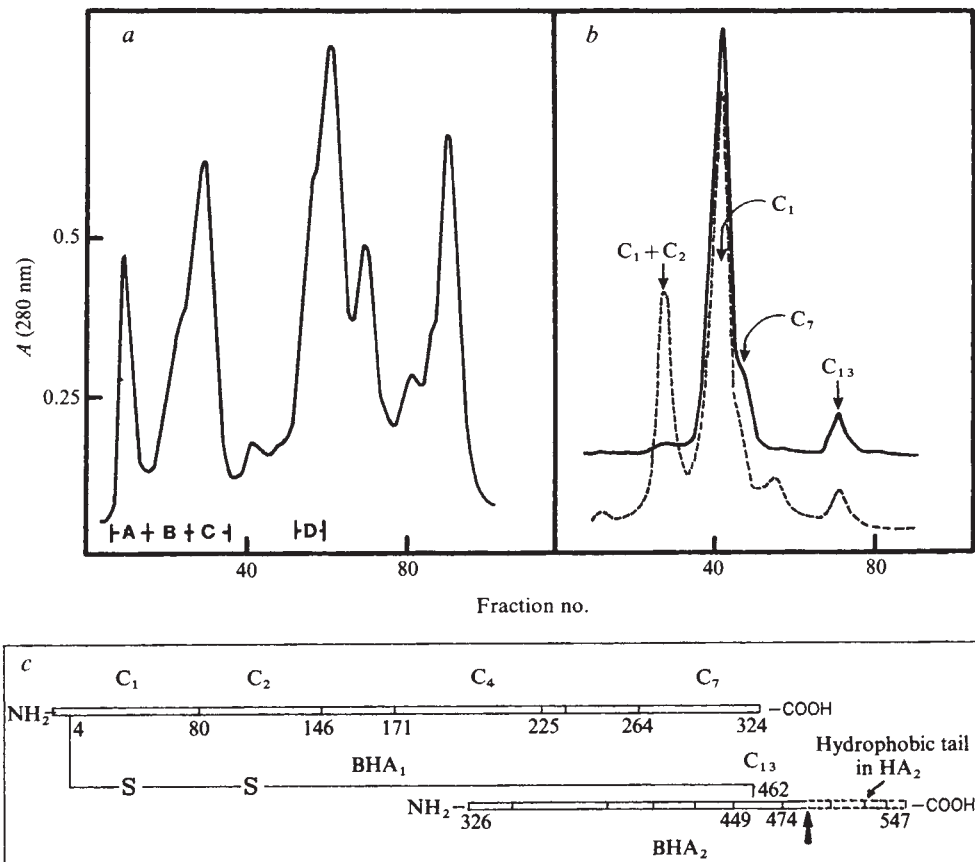
Table 1 Amino acid composition of disulphide-containing peptides isolated by HPLC from proteolytic digest of BHA cyanogen bromide fragments

	A	B	C	D	E	F	G	H
Cys	1.8 (2)	1.7 (2)	1.7 (2)	2.0 (2)	1.8 (2)	2.0 (2)	1.4 (2)	1.8 (2)
Asp	2.9 (3)	2.7 (2)	3.3 (3)	2.8 (2)	4.7 (4)	2.0 (2)	1.3 (1)	4.2 (4)
Threonine	—	4.6 (5)	0.2 (0)	4.6 (5)	0.8 (1)	—	1.1 (1)	—
Ser	1.6 (2)	1.3 (1)	0.8 (1)	0.6 (0)	1.0 (1)	—	—	2.6 (3)
Glu	1.8 (2)	3.5 (3)	1.0 (1)	2.6 (2)	2.9 (3)	1.2 (1)	2.0 (2)	2.5 (2)
Pro	1.2 (2)	4.1 (4)	1.0 (1)	4.3 (4)	—	—	—	2.2 (2)
Gly	2.6 (3)	2.5 (2)	3.0 (3)	3.1 (2)	3.2 (3)	—	1.4 (1)	2.8 (3)
Ala	0.9 (1)	1.2 (1)	1.0 (1)	1.3 (1)	1.2 (1)	—	—	1.8 (2)
Val	0.9 (1)	2.2 (2)	—	1.2 (1)	—	—	—	0.7 (1)
Met	—	—	—	0.1 (0)	—	† (1)	—	—
Ile	0.7 (1)	2.1 (2)	0.9 (1)	2.5 (2)	1.9 (2)	—	—	—
Leu	4.5 (5)	3.8 (3)	3.2 (3)	4.4 (3)	1.6 (1)	—	1.6 (2)	2.1 (2)
Tyr	—	0.6 (1)	—	—	1.2 (2)	—	—	1.6 (2)
Phe	—	1.0 (1)	—	1.1 (1)	1.7 (2)	—	—	2.9 (3)
His	—	2.0 (2)	—	2.0 (2)	1.9 (2)	—	—	—
Lys	—	3.2 (3)	—	2.0 (2)	1.1 (1)	—	1.0 (1)	1.0 (1)
Arg	1.0 (1)	—	0.9 (1)	0.1 (0)	0.3 (0)	—	—	1.0 (1)

Numbers in parentheses indicate predicted composition.

† Not quantifiable, present as oxidized homoserine lactone and homoserine.

Fig. 1 Fractionation of the CNBr fragments of BHA by gel permeation HPLC. *a*, Separation of peptides from intact BHA. *b*, Separation of peptides released from peak B (—) and peak C (---) after full reduction and alkylation. BHA (100 nmol) was prepared from virus grown in the allantoic cavity of eggs¹², cleaved with CNBr in 70% formic acid for 24 h and directly injected in 4 aliquots onto two Waters I-125 columns run in series using a Waters M6000A pump and M450 detector (Milford). Columns were run in 6 M urea containing 0.2 M formic acid at 0.5 ml min⁻¹ and monitored for absorbance at 280 nm; 0.5-ml fractions were collected and combined to give peaks A, B, C and D (*a*). Material in peaks B and C was desalted by dialysis or by absorption to Sep-Pak cartridges¹³, fully reduced and alkylated in 6 M guanidine at pH 8.5 (ref. 12) and directly re-injected onto the gel permeation columns (*b*). Fractions corresponding to absorbance peaks were pooled, urea removed, and amino acid compositions and amino-terminal sequences determined using methods described elsewhere^{7-9,12}. CNBr peptides C₁, C₂, C₇ and C₁₃ were identified. *c*, Line drawing of the location of sites of CNBr cleavage of HA and the numbering of the half-cystines in fragments referred to in the text. The arrow indicates the approximate site of cleavage by bromelain.



For the sake of clarity the original nomenclature⁷ has been changed. Thus, here the partial cleavage product (C₁ + C₂), and peptides C₁, C₂, C₄, C₇ and C₁₃ correspond to BHA₁-C₁, BHA₁-C₂, BHA₁-C₄, BHA₁-C₃ and BHA₂-C₄, respectively, in Waterfield *et al.*⁷.

(Fig. 1*b*). Appropriate fractions were pooled and the amino acid composition and amino-terminal sequence determined. Previous studies on the cyanogen bromide (CNBr) cleavage products of BHA₁ (refs 7, 8) and BHA₂ (ref. 9), knowledge of the complete amino acid sequence³ and amino-terminal sequence analysis using methods described elsewhere^{7,9} allowed us to establish the identity of the disulphide-linked fragments. Figure 1*c* shows the location of CNBr cleavages. Previous studies have shown that a partial CNBr cleavage occurs between residues 80 and 81 at a Met-Glu bond. This partial cleavage results in formation of a fragment containing C₁ joined to C₂. Peak B contained the partial cleavage fragment (C₁ + C₂), fragment C₁ (which results from complete cleavage), the C-terminal fragment C₇ of BHA₁ and fragment C₁₃ of BHA₂. Complete cleavage had occurred in the material eluted in peak C. After full reduction and re-chromatography, peak C was shown to contain peptides C₁ and C₇ of BHA₁ and C₁₃ of BHA₂ (Fig. 1*b*). Peak D was found to contain a mixture of fragments C₂ and C₄ of BHA₁. Fragment C₄ is known to contain no half-cysteine residues.

These results suggested that peptide C₂ of BHA₁ contained an intrachain disulphide bond and C₁ and C₇ of BHA₁ were linked to C₁₃ in BHA₂.

CNBr fragment B was digested with *Staphylococcus aureus* V8 protease and the whole digest fractionated by reverse-phase HPLC on a C8 column (Fig. 2*a*). Fractions corresponding to optical density peaks were pooled, aliquots oxidized with performic acid and amino acid analyses carried out to locate half-cysteine-containing peptides. A second portion of the V8 digest was further digested with trypsin and chromatographed and analysed in identical conditions (Fig. 2*b*). The amino acid compositions of pooled fractions containing half-cysteine are given in Table 1 and the peptides can be located in Fig. 3 by the residue numbers of the half-cysteine residues.

Two V8 peptides (A and B) which eluted at the end of the reverse-phase column run were identified by amino acid

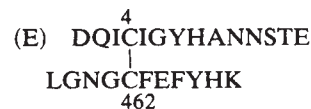
composition, using the single letter code of Dayhoff¹⁰, as



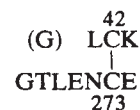
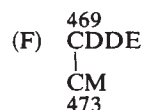
Further digestion with trypsin allowed identification of a tryptic cleavage fragment (C) which was derived from the V8 peptide (A). The V8 peptide (B) was recovered as peptide (D) after tryptic cleavage:



It was also possible to identify the following peptide (E)

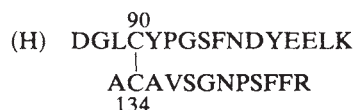


Re-chromatography of the digest using a flow rate of 0.5 ml min⁻¹ allowed resolution of the peptides eluted early and made it possible to identify (F) and (G):



These peptides allowed assignment of the single interchain bond which joins half-cystines 4 of BHA₁ and 462 of BHA₂ and also the assignment of intrachain disulphides between half-cystine residues 42 and 273, 55 and 67, 277 and 301 of BHA₁ and between 469 and 473 in BHA₂.

Assignment of the intrachain disulphide bond in CNBr peptide C₂ of BHA₁ was made by fractionation of a tryptic digest of a mixture of C₂ and C₄ which had been purified as peak D by gel permeation HPLC. The elution profile is shown in Fig. 2c and the composition of peptide (H), recovered in peak H, is given in Table 1. This compositional analysis confirmed the location of a disulphide bond between half-cystines 90 and 134:



The peptides isolated from BHA allow the assignment of disulphide bonds which link the twelve half-cystine residues of BHA. The results are summarized in Fig. 3. A similar study of the disulphide bonds of the haemagglutinin from Hong Kong virus A/Memphis/102/72 (H3N2) has been reported by

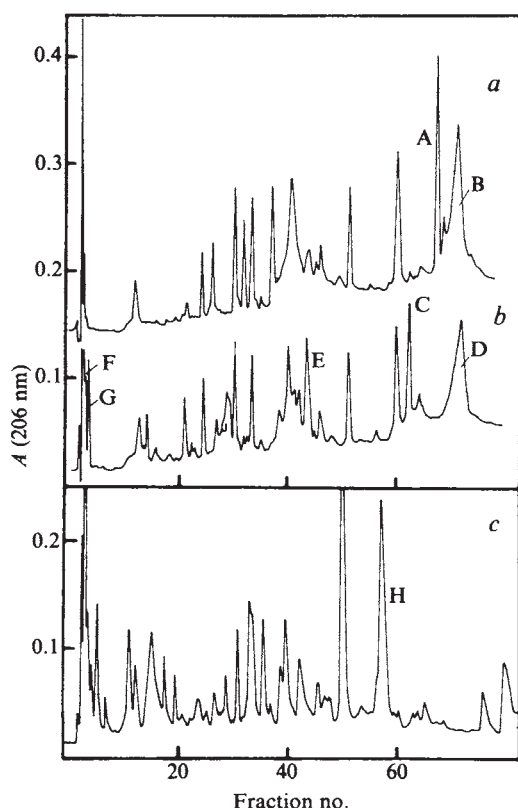


Fig. 2 Separation of enzymatic digests of disulphide-containing CNBr fragments by reverse-phase HPLC. Fragment C (purified as described in Fig. 1 legend) was digested with *S. aureus* V8 protease (70 nmol, enzyme/substrate 1:18) in 0.1 M ammonium bicarbonate buffer for 18 h and directly injected onto a Dupont Zorbax C8 column which was eluted with a gradient of acetonitrile in 10 mM ammonium acetate buffer, pH 6.5 (ref. 13). *a*, Column effluent was monitored for absorbance at 280 and 206 nm using two Uvicord S monitors (LKB Instruments) equipped with HPLC flow cells. Fractions (0.75 ml) were collected, peaks pooled by optical density and aliquots oxidized with performic acid and the amino acid composition determined. *b*, An aliquot of the V8 digest was further digested with trypsin (enzyme substrate 2% by weight) and analysed by the same methods. Material from pool D (see Fig. 1) was digested with 2% trypsin (by weight) and analysed by chromatography on Zorbax C8 as described for pool C. *c*, Separation of tryptic peptides. Peptides with disulphide bonds are indicated by letters referred to in the text.

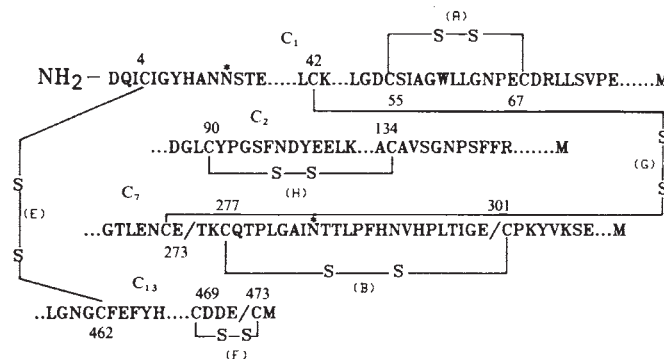


Fig. 3 The disulphide bonds of the bromelain-released fragments of the haemagglutinin of A/Japan 305/57 (H2N1). * Indicates the attachment site for carbohydrate side chains.

Dopheide and Ward¹¹ who present evidence that half-cystine 139 of HA₁ is linked to half-cystine 472 in HA₂ in the H3 strain and suggest that residue 97 may be linked to residue 476. Our results show that half-cystine residues homologous to those at positions 139, 97, 472 and 476 are arranged differently in the H2 virus HA. Because the position of half-cystine residues is highly conserved in haemagglutinins of viruses of both subtypes³ it is unlikely that the disulphide bonds differ. Analyses of the three-dimensional structure of the BHA fragment from an H3 virus (X-31) indicates a single interchain bond⁵ and it is likely that methodological difficulties remain to be resolved in the analysis of the disulphide-linked thermolytic peptides isolated by Dopheide and Ward¹¹. In agreement with their results, we assign an interchain bond between half-cystine residues 14 of HA₁ and 465 in HA₂ in the H3 strain, which are homologous with residues 4 of BHA₁ and 462 in BHA₂ in the H2 strain. Our results also agree with those of Dopheide and Ward¹¹ in the placement of intrachain disulphides between half-cystines at residues 52 and 277, 64 and 76, and 281 and 305 in HA₁ of the H3 strain, which are homologous to disulphides in HA₁ of the H2 strain which link residues 42 and 273, 55 and 67 and 277 and 301.

The results presented here establish disulphide bonds for the region of HA which is external to the membrane of the virus. A further four half-cystine residues are located near the C terminus of the HA of A/Japan 305/57 (H2N1)—two of these are conserved in the HAs of all strains examined (543 and 546)³. These half-cystines may be located inside the virus particle, below the lipid bilayer because N-terminal to these residues there is a 26–30-residue hydrophobic sequence that would be expected to cross the lipid bilayer. The remaining two half-cystines of this HA, which are not conserved in other strains, may simply serve as hydrophobic or uncharged amino acids. Whether these four half-cystines are disulphide bonded remains to be established. The contribution of the disulphide bonds to the three-dimensional structure of the HA is described by Wilson *et al.*⁶.

Received 30 October; accepted 23 December 1980.

- Porter, A. G. *et al.* *Nature* **282**, 471–477 (1979).
- MinJou, W. *et al.* *Cell* **19**, 683–696 (1980).
- Gething, M. J., Bye, J., Skehel, J. J. & Waterfield, M. D. *Nature* **287**, 301–306 (1980).
- Brand, C. M. & Skehel, J. J. *Nature new Biol.* **238**, 145–147 (1972).
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. in *Structures and Variations in Influenza Virus* (Developments in Cell Biology 5) (eds Laver, W. G. & Air, G.) 339–384 (Elsevier, Amsterdam, 1980).
- Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366–373 (1981).
- Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. *Br. med. Bull.* **35**, 57–64 (1979).
- Waterfield, M. D., Skehel, J. J., Nakashima, Y., Gurnett, A. & Bilham, T. in *The Influenza Virus Haemagglutinin* (Topics in Infectious Diseases) Vol. 3 (eds Laver, W. G., Bachmayer, H. & Weil, R.) 167–169 (Springer, Vienna, 1978).
- McCauley, J., Skehel, J. J. & Waterfield, M. D. in *The Influenza Virus Haemagglutinin* (Topics in Infectious Diseases) Vol. 3 (eds Laver, W. G., Bachmayer, H. & Weil, R.) 81–192 (Springer, Vienna, 1978).
- Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure* Vol. 5, D2 (National Biomedical Research Foundation, Maryland, 1972).
- Dopheide, T. A. & Ward, C. A. *FEBS Lett.* **110**, 181–182 (1980).
- Skehel, J. J. & Waterfield, M. D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 93–97 (1975).
- Waterfield, M. D. & Scrase, G. T. in *Biological/Biomedical Applications of Liquid Chromatography III* (ed. Hawk, G.) (in the press).

BOOK REVIEWS

Swimming against the tide

Eric Ashby

The Politics of Self-Sufficiency. By Michael Allaby and Peter Bunyard. Pp.242. (Oxford University Press: 1980.) Hbk £7.95, \$24.95; pbk £3.95.

IF YOU see someone swimming against the tide, watch him. He may sink without trace; he may bring ashore a new idea. It is in this spirit that I commend this duet by Michael Allaby and Peter Bunyard. A duet? Yes, for in each chapter first one author, then the other, has his say. The two authors agree in principle that self-sufficiency is a Good Thing. They do not agree about the way in which self-sufficiency might be achieved nor about the part it might play in Western industrial society.

On the first page the authors define what they are writing about. It is not the self-sufficiency of pioneers colonizing a hostile environment, often by a ruthless exploitation of nature; it is self-sufficiency as an alternative "lifestyle": the community which grows its own food (preferably without artificial manures), generates its own energy (or some of it) from sunlight and the burning of timber, disposes of its own wastes, imports such things as it cannot make in exchange for its own surplus, by barter, not money. There are already little communities of this kind in Britain, and the Israeli kibbutz and Chinese commune are viable versions of a similar idea. Indeed, if one is looking for historical precedents for self-sufficiency, the mediaeval monastery is surely an impressive example.

The self-sufficient community, like the monastic community, can be sustained as a minority way of life in an industrial pluralistic democracy. It is when enthusiasts for self-sufficiency urge us to make it a majority way of life that doubts arise. All sorts of awkward questions have to be answered. Is it, in any case, a "better" lifestyle? Could such a lifestyle be sustained by the minorities who enjoy it if they were not embedded in a conventional capitalist society which can come to their rescue when they need to travel, to listen to music, to get medical attention, to type and publish books such as this one, to replace tools? And, since this is a book about the politics of self-sufficiency, under what kind of political system could self-sufficiency ever become a majority lifestyle?

The virtue of the book is that it does examine these questions honestly. It considers the prospects for a mass-movement toward self-sufficiency. The

prophets have to step warily between the ideologies of anarchy on one side and fascism on the other. Except for hermits, self-sufficiency implies co-operation, and "the moment a group of people act collectively they cease to act anarchically". So the self-sufficient society would have to submit itself to some sort of government. If it is to be a government which rejects an economic system based on motives of profit, then "the economy itself would be changed from a market to a command type. The injunction to industrialists would no longer be the simple 'Be profitable!' but the much more complex 'Make this, but not that'." The only known patterns of government capable of behaving this way are far-left or far-right dictatorships. The authors don't relish either of these. They seek a middle way: some communal ownership of property; some concession to the fact (which they freely admit) that an atavistic retreat to a bucolic past is simply not on; some compromise with the harsh reality of the human condition, which is that greed, selfishness and love of comfort are majority attitudes, and that beneficence, self-denial and austerity are likely to remain minority attitudes. They adduce plenty of evidence that small communities can — and some do — drastically reduce their reliance upon the products of industrial society, and consume much less per capita than the average consumption of the limited resources of nature. They make the point clearly that they do not consider self-sufficiency to be a middle class élitist indulgence or a highbrow name for drop-outs. Self-sufficiency is a deep — they call

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Self-sufficiency in the Middle Ages — how viable today?

it a "spiritual" — protest against a way of life which is carrying us out on a tidal current to God knows what dangers.

Anyone who seriously contemplates self-sufficiency in political terms had better stand at Waterloo Station in the rush hour and ask himself exactly how any political movement is going to lead this surge of commuters into such a different lifestyle, and ensure that they are adequately fed, sheltered and kept warm in (say) the idyllic conditions of rural Devon and Cornwall. (And, incidentally, ask himself what would happen to Devon and Cornwall if such a miracle occurred.) But this would be a slick and insensitive way to end this review: for there is one important conclusion to be drawn from this book, though it may not be a conclusion the authors approve. It is that little self-sufficient communities such as Allaby and Bunyard describe have an important part to play in contemporary Britain. Like monasteries in the Dark Ages, they affirm that there is another, and perhaps a better, way to live than the way most people live and that it is not out of reach. They are not composed of cranks. They are men and women who believe in swimming against the tide. They may sink without trace. But, like monasteries in the Dark Ages, they may carry the seeds for a new kind of civilization, one able to transform the taint of unemployment into the boon of leisure for those without work, and the drudgery of attending an assembly line for the satisfaction of manual labour for those who have work. □

Lord Ashby is Chancellor of Queen's University, Belfast, and a Fellow of Clare College, Cambridge.

Health risk: just a guessing game?

T. A. Connors

Ethylene Dichloride: A Potential Health Risk? 5th Banbury Report. Edited by B. Ames, P. Infante and R. Reitz. Pp.350. (Cold Spring Harbor Laboratory: 1980.) \$45, \$54 outside USA.

"WELL, is ethylene dichloride a health risk?" is the question that comes to mind after reading this book. The comment does not imply any specific criticism of the report but rather sums up our present knowledge of the toxicology of a wide variety of chemicals to which human beings are exposed. We suspect from animal and other laboratory tests that many chemicals may be a risk to human health as carcinogens, teratogens, mutagens and so on, but we do not have reliable methods to determine whether there is an actual risk in any situation or what the degree of that risk may be.

Ethylene dichloride is a very important chemical indeed, the US production in 1977 amounting to 11,000 million pounds in weight. It is used mainly as an intermediate, so most high-level exposure is in the chemical industry, but significant amounts are also used in the textile and food industries and in agriculture, for example. The possibility that ethylene dichloride might be carcinogenic was first suggested when it was shown to be mutagenic in bacteria. In a large international study, tests of this type have been shown — at least for certain classes of chemical — to correlate reasonably well with animal carcinogenicity tests which in turn have given positive results for most human carcinogens identified by epidemiological or case control studies. This Banbury Report summarizes what is known of the toxicology of ethylene dichloride and one learns that it can be metabolized *in vivo* to an alkylating agent which can react with DNA and initiate the events that many people think can lead to cancer. One also learns that other short-term tests, using *Drosophila* and yeast, for example, have given positive results and that the chemical, like many of its congeners, is carcinogenic in rats and mice. However not all of the data are unequivocal. In a well-controlled study, Maltoni and his colleagues could find no evidence for the carcinogenicity of ethylene dichloride in rats and they question the validity of earlier studies from a number of aspects including "the professionalism of the team carrying on treatment, control of animals and autopsies". Furthermore, no teratogenicity or major reproductive toxicity was found either in animals or in exposed workers, nor was there any evidence of an increased incidence of cancer in workers exposed to ethylene dichloride for 20 years or more.

All in all the book is a wonderful

miniature of the multidisciplinary field of toxicology and the problems that arise in interpreting data from different types of test, in choosing between conflicting results and in attempting to assess the risk to human beings. Although scientists, as in this report, are to be commended for attempting to measure carcinogenic potency and to quantify risk, the equation contains so many unknowns that the assessed level of risk must be seen to be based on a working hypothesis with many assumptions rather than as the result of an extrapolation based on scientifically reliable methods. It follows then that in making the decision to restrict a chemical, to replace it with "safer" alternatives or to ban it altogether, economic and social argument must be considered as well as the scientific aspects. There is concern that a number of chemicals which have been used for many years and considered to be safe because there was no acute toxicity

associated with exposure, may nevertheless be highly dangerous in the long term. The Banbury Report, with contributions from industrialists, government scientists and university researchers, shows how widespread this concern is, but until animal experiments can be related to human situations the debate on the dangers of chemicals will continue.

However, there has already been one welcome outcome. Workers are nowadays exposed to less dirty, dusty and smelly conditions than even a few years ago, while urban environments are certainly more wholesome than they once were. Whatever the effects may be on the cancer incidence only time will tell; but for those people who believe that quality of life is at least as important as quantity of life then there is already a vast improvement. □

T.A. Connors is at the MRC Toxicology Unit, Carshalton, Surrey.

Catastrophes for starters

Colin Upstill

An Introduction to Catastrophe Theory. By P.T. Saunders. Pp.144. (Cambridge University Press: 1980.) Hbk £9.50, \$27.50; pbk £3.25, \$8.95.

ALL previous introductions to catastrophe theory have been either popularizations, devoid of any serious mathematics and often riddled with falsehoods, or works of considerable mathematical sophistication. This book succeeds in filling the gap between these extremes. It divides naturally into two parts. The first is an exposition of the theory, which deals with the singularities of smooth real valued functions, using only a limited mathematical vocabulary yet not glossing over anything of importance: structural stability, equivalence, the splitting lemma, determinism and universal unfoldings are all explained as clearly as one could wish. The second half of the book gives the flavour of applications of the theory by way of a selection of examples from the physical, social and biological sciences, all of which are to be found discussed in much greater detail elsewhere in the literature. The author is careful not to exaggerate the status of catastrophe theoretic models in the social and biological sciences, and concludes with a lucid discussion of the explanatory powers of the theory and some cautionary words on appropriate standards of judgement of applications in different disciplines.

My enthusiasm is not unqualified,

however. Central to catastrophe theory are the geometries of the catastrophes themselves, so it is inexcusable that their illustration in this book leaves something to be desired. Most of the relevant figures are slavishly and uncritically copied, with due acknowledgement, from the book by Bröcker and Lander (*Differentiable Germs and Catastrophes*; Cambridge University Press, 1975); some include labelling in a nomenclature incompatible with that used in Saunders's text, others include construction lines which are here irrelevant and unexplained, or fail to display the nature of the singularity at the origin of control space. Where the drawings are original, similar lack of care is evident. At the hyperbolic umbilic singularity, the rib (cusped edge) and the fold touch parabolically, yet here they are depicted as intersecting straight lines; the sketch of the lips event is quite dreadful.

The text does not suffer from any such sloppiness, but there are some omissions which are surprising in such a recent publication. Most notable is the absence of any reference to the work of the Russian mathematician V.I. Arnol'd, who is responsible for the enormous extension of Thom's classification to catastrophes of codimension > 4 , and for a rational system of symbols to label the catastrophes, which Saunders eschews — the Thomist "pet names" are fine if one is only considering a few low-dimensional catastrophes, but things get out of hand as the dimensionality and number of catastrophes in the list

increases. Explanations of the role of higher dimensional catastrophes as organizing centres and of the importance of compact catastrophes in applications would not come amiss; neither would some discussion of (or at least references to) recent vociferous criticisms of some applications of catastrophe theory in the social and biological sciences.

In a rapidly expanding subject it is inevitable that any book will be out of date as soon as it is published. However, much of this book could have been written four years ago. Whilst there is no claim to define

the "state of the art", it is an uneven treatment, to say the least, that for some applications covers material from one or two expository works of several years ago, but for others refers to a quantity of very specialized material of recent date. One is forced to conclude that the author has made no great effort to acquaint himself with recent work outside his own particular interests, which is a pity, because as a result he has written a good book when he could have written a superb one.

Colin Upstill is a Research Fellow at the H. H. Wills Physics Laboratory, University of Bristol.

Potpourri of nuclear structure theory

J. M. Irvine

Nuclear Spectroscopy. Lecture Notes in Physics, 119. Edited by G. F. Bertsch and D. Kurath. Pp.250. (Springer-Verlag: 1980.) DM33, \$19.50.

THE Gull Lake workshop on nuclear spectroscopy, held in the summer of 1979, reviewed most of the basic topics in theoretical nuclear structure physics. This book is a record of the deliberations of the participants.

The opening chapter by Gerry Brown takes us through the familiar analysis of the meson exchange nature of nuclear forces, linking it to the quark structure of the nucleon in the "little bag model" of Brown and Rho. The Brueckner-Bethe and Landau-Migdal theories of effective interactions are developed and the predictions for the bulk properties of nuclear matter presented. In other interesting papers, Dieter Kurath presents an introduction to the nuclear shell-model which is novel in that it includes a short discussion of the current shell-model computer codes, and George Bertsch discusses nuclear vibrations — both in the random phase approximation and Landau formalisms — and presents an up-to-date account of the new giant resonances.

Amand Faessler gives a unified account of deformed and transitional nuclei. Concentrating first on high spin states in deformed nuclei, he discusses the "back-bending" phenomena and the nature of Yrast traps, and goes on to give an account of the breakdown of pairing due to the Coriolis force in transitional nuclei. Franco Iachello describes the use of group theory in the analysis of nuclear spectroscopy with a natural emphasis on the successes of the interacting boson approximation.

The use of statistical analysis in describing complex spectroscopies has been advocated for a long time. At last there are signs that it is being put to use in nuclear physics and it is appropriate that the field should here be reviewed by J. B. French. Finally, there is an interesting appendix by Bertsch, Zamick and Mekjian which highlights some of the unsolved theoretical puzzles which abound in nuclear structure physics.

I found this an exceptionally readable account of current topics in nuclear theory. The level of presentation of the material is remarkably uniform for a multi-author volume. I am sure that students and research workers in the field will find this a useful addition to their bookshelves.

J. M. Irvine is Reader in Theoretical Physics at The University of Manchester.

where, although the relationships may be equally close and complex, they can be non-obligate and accidental, as with dipterous flies and many bacterial plant pathogens.

The book clearly points out the wide gaps in our knowledge and understanding of this subject. Many chapters end with questions for the reader or indicate need for further study. One contribution is even devoted to the search for the vector of a disease (lethal yellowing of coconut palm) where the very nature of the pathogen is still uncertain, although the reader may assume it to be a mycoplasma-like organism (MLO); a cixiid bug appears to be the most likely candidate as the vector.

The 13 chapters are written by different specialists, or groups of specialists who were clearly given a free hand and who have very different styles of writing. This variation in style adds to the interest but perhaps the editors might have tied the volume together a little more closely.

This thought-provoking book contains many references and will be of value and interest to researchers as well as to students and teachers. For example, how many of us have considered the possibility of a link between the use of mustard gas in the First World War and the recent spread of Dutch elm disease? The book suggests one. □

G. D. Heathcote is at Broom's Barn Experimental Station, Higham, Bury St Edmunds, Suffolk, and works on transmission of viruses by aphids.

Seabird synthesis

J.B. Nelson

Behavior of Marine Animals. Vol.4, Marine Birds. Edited by J. Burger, B.L. Olla and H.E. Winn. Pp.515. (Plenum: 1980.) £28.35, \$54.

STUDENTS of seabirds will warmly welcome this fine and wide-ranging compilation of data, ideas and references. Many important aspects of breeding biology such as choice of habitat, mate-selection, various factors and strategies relating to breeding success and the nature of communication behaviour are excellently reviewed. These articles form the core of the book but three others (seabirds at sea, the distribution of gulls in North America and, somewhat oddly placed, chemoreception in seabirds) stand outside this major theme.

Systematic use of headings and sub-headings produces welcome internal consistency within the volume but, in view of inevitable omissions in such a wide-ranging and essentially review work (references average 148 per article, excluding that by Southern), space could have been saved by editorial restrictions on the text authors devote to their own research. As it is, several contributors suddenly abandon the

Various vectors and plant diseases

G. D. Heathcote

Vectors of Plant Pathogens. Edited by Kerry F. Harris and Karl Maramorosch. Pp.467. (Academic: 1980.) £27, \$48.

THIS book succeeds admirably in what it sets out to do, providing much detailed, up-to-date information on the inter-relationships of many plant pathogens, their plant hosts and various vectors. It is

third in a series of four volumes being published on this general topic. The earlier volumes were concerned with aphids and leaf-hoppers, and, although the first chapter of this volume is devoted to pathogen-vector relationships within these groups of insects, most of the current book is concerned with other, generally less well-known vectors. These range from mealybugs, to fungi and to nematodes

review style and launch into detailed and specialized exposition. Undoubtedly, though, seabird workers will warmly welcome this work as evidence that their research, despite its heterogeneous and incomplete nature, nevertheless engenders valuable syntheses. Seabird studies, as one might expect after two highly productive decades, are moving away from general data gathering and towards more detailed investigations of specific problems. These contributions point the way to many such areas. Perhaps the chief merit of the book lies in its discussion of the relevance of this mass of work to well-recognized problems, and in the clarification of these problems, rather than in the formulation of new concepts. Old terms, hitherto loosely and confusingly employed — for instance synchrony, colony, fledging, starvation period, display, social stimulation and the Darling effect — are carefully dissected. Again, nowhere else in all the seabird literature can one find a properly detailed discussion of the effect of age on breeding success. As an example of "properly detailed" there is Burger's Table II, which gives information individually for 86 species under 10 subject headings.

Yet there are fundamental concepts, affecting most of these subjects, which receive no critical attention. I would have welcomed a thorough review of the concept of direct, density-dependent competition which I am convinced often does not operate among seabirds. Within the constraints of this volume, no single contributor could have tackled the subject of the regulation of seabird populations, but there are reliable and reviewable field data which bear directly on the validity of density-dependent regulation in seabirds.

Among the many ideas canvassed, I liked Hunt's emphasis on the need to distinguish a lifetime's reproductive output from that of any one year, an important but rather inaccessible statistic when considering the reproductive strategies so beloved of today's theoreticians. Ryder's excellent summary of age-effects brings out the intriguing phenomenon of reduced success in old individuals. In related vein, Montevecchi and Porter indicate the minimum detail which must be obtained on time and energy budgets before one can begin to theorize about optimal use of resources. They emphasize, as does Gochfield, what I too have suggested, that for many seabirds, especially males, the costliest part of the reproductive cycle is not chick-rearing but the pre-laying stage when sites are established (or re-

established) and courtship performed. Beer, bristling with ideas, emphasizes the importance of context on the form and function of display — a minefield of potential ambiguity. His article should be required reading for behaviourists, to steer them away from beguilingly simplistic approaches. But even he founders in the complexities inherent in the notion of "intention" in animals.

Gochfield must speak for many in his plea for adequate presentation of data (and he gives an excellent account of methods) to allow others to re-work them. Highly processed results may satisfy statisticians but are useless, except in a take-it-or-leave-it form, to the Indians.

Bridge-building is an important function of review articles and Brown diligently and perceptively connects the many scattered

observations on birds at sea with some oceanographic data. Even so, the key words remain "broad outline" in his statement that "we now know in broad outline the pelagic distribution of the common seabirds in summer at least, in almost all the world's oceans". Many of his valuable references will be new to most ornithologists.

Throughout the book, and rightly so, functional interpretations are stressed, for it is essential to generate testable ideas about the adaptive significance of the intricate web of behaviour by means of which animals cope with environmental demands. □

J.B. Nelson is Reader in Zoology at the University of Aberdeen, and the author of Seabirds: Their Biology and Ecology, published earlier this year by Hamlyn.

Mass spectrometry for biochemists

Howard R. Morris

Biochemical Applications of Mass Spectrometry: First Supplementary Volume. Edited by George R. Waller and Otis C. Dermer. Pp.1,279. (Wiley-Interscience: 1980.) £79.80, \$187.50.

THIS is the first supplementary volume of a previous work of the same title, and is a compendium of contributions by well-known authors in specialist fields. The chapters have been written at a level suitable for postgraduate students, and in general they provide useful background reading for the "layman", whether he be a newcomer to research in one of these fields, or a biochemist or biologist seeking a solution to a particular problem.

The editors predict in the preface that certain sections are unlikely to be comprehensive in content, and this is indeed so in a number of cases. For example, in the section on instrumentation there is, inexcusably, no reference to one of the most significant and highly relevant developments in the area of biochemical mass spectrometry, the high-field magnet. The new range of commercial instruments based upon this concept (giving high sensitivity operation up to m/e 1800 — "extended" — or m/e 3000 — "high-field") is of special interest and importance to the biochemist, since he is often dealing with small amounts of high-molecular-weight polar compounds.

A considerable portion of this large book is taken up with matters which are not *per se* biochemical mass spectrometry, such as basic principles, and a listing of the equipment in several different spectrometry laboratories. The choice of these laboratories would appear to be somewhat arbitrary, since a number of them are not noted for major contributions

in the field of biochemistry. The validity of detailed descriptions of mass spectrometric data systems is in any event debatable for the intended audience. Similarly, in the section on interpretation of mass spectra, a series of methods and computerized techniques is given which, as far as one can tell, have not yet been applied to the solution of biochemical problems, and would therefore more properly have appeared in a subsequent volume.

In contrast, the section covering applications makes very worthwhile reading, since this is the real meat of biochemical mass spectrometry. The chapter on antibiotics is excellent, and that on proteins is commendably comprehensive, despite a few inevitable inaccuracies in the interpretation of some of the details of the work covered. Unfortunately, not all of the authors are so thorough; for example the section on tetrapyrroles is notable for its superficial coverage. Finally, the descriptions of some interesting new applications make enjoyable reading, in particular the extension of mass spectrometric techniques into the study of volatiles emitted by human beings; one might predict a rapid growth of interest in this area since the information provided would appear to be highly relevant to the physiological status of the body in both health and disease.

As a whole, *Biochemical Applications of Mass Spectrometry* is a useful and often stimulating reference work, providing the biochemical student or researcher with a basic background from which to make a more detailed exploration of the literature. □

Howard R. Morris is Professor of Biological Chemistry in the Department of Biochemistry at Imperial College, University of London.

The second of the six core volumes of the *Handbook of Applicable Mathematics* has been published by John Wiley (for review see *Nature* 288, 37). Volume II, *Probability*, edited by Emlyn Lloyd, is available as a single volume at £32.50, \$85. Subscription price for the core volumes is £27.50, \$75.

BOOKS RECEIVED

Mathematics

- ANDERSON, S. *et al.* Statistical Methods for Comparative Studies. Techniques for Bias Reduction. Pp.289. ISBN 0-471-04838-0. (Wiley: 1980.) £13.50.
- GAL, S. Search Games: Mathematics in Science and Engineering, Vol. 149. Pp.216. ISBN 0-12-273850-0. (Academic: 1980.) \$20.
- STEWART, W. E., RAY, W. H. and CONLEY, C. C. (eds). Dynamics and Modelling of Reacting Systems. Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, The University of Wisconsin-Madison, October 1979. Pp.413. ISBN 0-12-669550-4. (Academic: 1980.) \$27.50.

Astronomy

- BALIAN, R., AUDOZE, J. and SCHRAMM, D. N. (eds). Physical Cosmology. Proceedings of Les Houches Summer School, Session XXXII, July 1979. Pp.668. ISBN 0-444-85433-9. (North-Holland: 1980.) \$109.75, Dfl. 191.
- HALLIDAY, I. and MCINTOSH, B. A. (eds). Solid Particles in the Solar System. Symposium No. 90 Organized by the IAU in co-operation with COSPAR held at Ottawa, August 1979. Pp.432. Hbk ISBN 90-277-1164-X; pbk ISBN 90-277-1165-8. (Reidel: 1980.) Hbk Dfl. 95, \$49.95; pbk Dfl. 50, \$26.50.

Physics

- BERTOZZI, W., COSTA, S. and SCHAEFER, C. (eds). Electron and Pion Interactions with Nuclei at Intermediate Energies. Studies in High Energy Physics, Vol.2. Proceedings of the International School sponsored by the University of Rome, held at the Centre Studi in Ariccia, Rome, June 1979. Pp.703. ISBN 3-7186-0015-3. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) \$55, Dfl.108.
- BERTRAND, F. E. (ed.). Giant Multipole Resonances. Nuclear Science Research Conference Series, Vol.1. Proceedings of the Conference held at the American Museum of Science and Energy, Oak Ridge, Tennessee, October 1979. Pp.482. Flexi ISBN 3-7186-0029-5. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) \$42.50, Dfl.83.
- GIBSON, W. M. and POLLARD, B. R. Symmetry Principles in Elementary Particle Physics. Pp.389. Hbk ISBN 0-521-20787-8; pbk ISBN 0-521-29964-0. (Cambridge University Press: 1980.) Hbk np; pbk £9.95.
- HARGRAVES, R. B. (ed.). Physics of Magnetic Processes. Pp.585. Hbk ISBN 0-691-08259-6; pbk ISBN 0-691-08261-8. (Princeton University Press: 1980.) Hbk £21.90; pbk £8.40.
- HOOF, G. *et al.* Recent Developments in Gauge Theories. NATO Advanced Study Series. Series B: Physics Vol.59. Pp.438. ISBN 0-306-40479-6. (Plenum: 1980.) \$49.50.
- HOROWITZ, P. and HILL, W. The Art of Electronics. Pp.716. Hbk ISBN 0-521-23151-5; pbk ISBN 0-521-29837-7. (Cambridge University Press: 1980.) Hbk £35; pbk £12.50.
- JERRARD, H. G. Electro-Optics/Laser International '80 UK Conference Proceedings, Brighton, March 1980. Pp.401. ISBN 0-86103-035-4. (IPC Science and Technology Press Ltd, Guildford, 1980.) £22.50.
- KOLLMANN, W. Computational Fluid Dynamics, Vol.2. Pp.266. ISBN 0-89116-192-9. (Hemisphere Publishing: 1980.) \$55.
- PEEBLES, P. J. E. The Large-Scale Structures of the Universe. Pp.421. Hbk ISBN 08240-5. (Princeton University Press: 1980.) Hbk £16.50; pbk £5.50.
- SCHUTZ, B. F. Geometrical Methods of Mathematical Physics. Pp. 250. Hbk ISBN 0-521-23271-6; pbk ISBN 0-521-29887-3. (Cambridge University Press: 1980.) Hbk £18; pbk £7.50.
- TAMURA, T., NATOWITZ, J. B. and YOUNGBLOOD, D. H. (eds). Continuum Spectra of Heavy Ion Reactions. Nuclear Science Research Conference Series, Vol.2. Proceedings of the International Symposium, San Antonio, Texas, December 1979. Pp.476. Flexi ISBN 3-7186-0028-5. (Harwood Academic Publishers: P.O. Box 786, Cooper Station, New York 10003, 1980.) \$48.50, Dfl. 96.

Chemistry

- CARR, P. W. and BOWERS, L. D. Immobilized Enzymes in Analytical and Clinical Chemistry. Fundamentals and Applications. Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications, Vol.56. Pp.460. ISBN 0-471-04919-0. (Wiley: 1980.) £24.
- CHAMOT, E. M. and MASON, C. W. Handbook of Chemical Microscopy, Vol. II. Chemical Methods and Inorganic Qualitative Analysis. 2nd Edn. Pp.438. ISBN 0-471-04122-X. (Wiley: 1980.) £14.75.
- DE MAYO, P. (ed.). Rearrangements in Ground and Excited States, Vol. 3: Organic Chemistry, Vol. 42. Pp.607. ISBN 0-12481303-8. (Academic: 1980.) Np.
- HEAL, H. G. The Inorganic Heterocyclic Chemistry of Sulphur, Nitrogen and Phosphorus. Pp.271. ISBN 0-12-335680-6. (Academic: 1980.) £35.80, \$82.50.
- HUTZINGER, O. (ed.). The Handbook of Environmental Chemistry, Vol.1. Part A. The Natural Environment and the Biogeochemical Cycles. Pp.258. ISBN 3-540-09688-4. (Springer: 1980.) DM 98, \$57.70.
- JOHNSON, B. F. G. (ed.). Transition Metal Clusters. Pp.681. ISBN 0-471-27817-3. (Wiley: 1980.) Np.
- KRYT, D. (ed.). Dictionary of Chemical Terminology. In five languages: English, German, French, Polish and Russian. Pp.612. ISBN 0-4444-99788-1. (Elsevier Scientific: 1980.) \$95, Dfl.195.

- WEBB, G. A. (Sen. Reporter) Nuclear Magnetic Resonance. Vol.9: A Review of the Literature published between June 1978 and May 1979. Pp.336. ISBN 0-85186-960-2. (The Royal Society of Chemistry: London, 1980.) £48, \$129.

Technology

- BARD, A. J. and FAULKNER, L. R. Electrochemical Methods. Fundamentals and Applications. Pp.718. ISBN 0-471-05542-5. (Wiley: 1980.) £14.70.
- BASELT, R. C. Biological Monitoring Methods for Industrial Chemicals. Pp.301. ISBN 0-931890-04-7. (Wiley: 1980.) £24. USA: Biomedical Publications, PO Box 368, Canton, Connecticut 06019.
- BRITISH ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE. Energy in the Balance. Some Papers from BA79 given at the Annual Meeting of the BAAS Heriot-Watt University, Edinburgh, 1979. Pp.234. Flexi ISBN 0-86103-031-1. (Westbury House: PO Box 63, Bury Street, Guildford, Surrey GU2 5BH, 1980.) Np.
- DUFFIE, J. A. and BECKMAN, W. A. Solar Engineering of Thermal Processes. Pp.762. ISBN 0-471-05066-0. (Wiley: 1980.) £13.80.
- GROSS, W. A. *et al.* Fluid Film Lubrication. Pp.774. ISBN 0-471-08357-7. (Wiley: 1980.) £18.65.
- HUGHES, D. A Literature Survey and Design Study of Fume Cupboards and Fume-Dispersal Systems. Occupational Hygiene Monograph No. 4, 1980. Pp.82. Flexi ISBN 0-905927-50-8. (Science Reviews Ltd: 1980.) £7.50.
- KAVANAGH, R. (ed.). Energy Systems Analysis. Commission of the European Communities. Proceedings of the International Conference held in Dublin, October 1979. Pp.678. ISBN 90-277-1111-9. (Reidel: 1980.) Dfl. 90, \$47.50.
- McAULIFFE, C. A. Hydrogen and Energy. Pp.109. ISBN 0-333-18432-7. (Macmillan: 1980.) £12.
- MURR, L. E. (ed.). Solar Materials Science. Based on the Distinguished Lecture Series Sponsored by the New Mexico Joint Center for Materials Science, September-December 1979. Pp.788. ISBN 0-12-511160-6. (Academic: 1980.) \$35.
- ROBINSON, F. A. (ed.). Environmental Effects of Utilising More Coal. Special Publication No. 37. Proceedings of a Conference organised by The Council for Environmental Science and Engineering. Pp.204. Flexi ISBN 0-85186-805-3. (Royal Society of Chemistry: Burlington House, London, W1V 0BN, 1980.) £9.50.

Earth Sciences

- BALLANCE, P. F. and READING, H. G. Sedimentation in Oblique-slip Mobile Zones. Special Publication No.4 of the International Association of Sedimentologists. Pp.266. Flexi ISBN 0-632-00607-2. (Blackwell Scientific: 1980.) £14.50.
- CARRERAS, J. *et al.* (eds). Shear Zones in Rocks. Journal of Structural Geology Vol.2, Number 1/2. Pp.287. Flexi ISBN 0-08-026244-9. (Pergamon: 1980.) £15, \$35.
- DERRY, D. R. World of Geology and Mineral Deposits. Pp.110. ISBN 0-470-26996-0. (Wiley: 1980) £20.
- GREESON, P. E., CLARK, J. R. and CLARK, J. E. (eds). Wetlands Functions and Values: The State of our Understanding. Proceedings of the National Symposium, Disneyworld Village, Lake Buena Vista, Florida, November 1978. Pp.674. (American Water Resources Association: St. Anthony Falls Hydraulic Laboratory, Mississippi River at 3rd Avenue, S.E. Minneapolis, Minnesota 55414, 1980) \$49 + postage outside US and its territories.
- JONES, D. K. C. (ed.) The Shaping of Southern England. Institute of British Geographers Special Publication, No.11. Pp.274. ISBN 0-12-388950-2. (Academic: 1980.) £14.60, \$34.
- MESSEL, H. *et al.* Surveys of Tidal River Systems in the Northern Territory of Australia and their Crocodile Populations. Monograph 9: Tidal Waterways of Castlereagh Bay and Hutchinson and Cadell Straits. Pp.132. Flexi ISBN 0-08-024801-2. (Pergamon: 1980.) £11.
- SCHUMACHER, M. M. (ed.). Enhanced Recovery of Residual and Heavy Oils. 2nd Edn. Pp.378. ISBN 0-8155-0816-6. (Noyes Data Corporation: Park Ridge, New Jersey, U.S.A., 1980.) \$48.
- TURNER, F. J. Metamorphic Petrology: Mineralogical, Field, and Tectonic Aspects. 2nd Edn. Pp.524. ISBN 0-07-065501-4. (Hemisphere Publishing: 1981.) Np.

Biological Sciences

- ALEXANDER, M. (ed.). Advances in Microbial Ecology, Vol.4. Pp.247. ISBN 0-306-40493-1. (Plenum: 1980.) \$32.50.
- APPLETON, D. R., SUNTER, J. P. and WATSON, A. J. (eds). Cell Proliferation in the Gastrointestinal Tract. Pp.428. ISBN 0-272-79597-6. (Pitman Medical: 1980.) £25.
- BAKER, R. (ed.). Controlled Release of Bioactive Materials. Based on the Symposium at the 6th International Meeting in New Orleans, Louisiana, August 1979. Pp.473. ISBN 0-12-074450-3. (Academic: 1980.) \$34.50.
- BRITISH MUSEUM (NATURAL HISTORY) Man's Place in Evolution. Pp. 108. Hbk ISBN 0-521-23177-9; pbk ISBN 0-521-29849-0. (Cambridge University Press: 1980.) Hbk £12; pbk £3.95.
- CARR, I. and DAEMS, W. T. (eds). The Reticuloendothelial System. A Comprehensive Treatise. Vol.1, Morphology. Pp.793. ISBN 0-306-40291-1. (Plenum: 1980.) \$49.50.

- COLD SPRING HARBOR SYMPOSIA ON QUANTITATIVE BIOLOGY. Vol. XLIV: Viral Oncogenes. Parts 1 & 2. Pp.1322. ISBN 0-87969-043-6. (Cold Spring Harbor Laboratory: 1980.) Np.
- DORNFELD, E. J. The Butterflies of Oregon. Pp.276. Flexi ISBN 0-917304-58-6. (Timber Press: P.O. Box 92, Forest Grove, Oregon 97116, 1980.) \$24.95.
- DURAND, J. R. and LEVEQUE, C. Flore et Faune Aquatiques de l'Afrique. Sahelo-Soudanienne, Tome I. Collection Initiations-Documents Techniques No.44. Pp.389. ISBN 2-7099-0521-3. (Editions de l'Office de la Recherche Scientifique et Techniques Outre-Mer: Paris, 1980.) Np.
- FINN, C. A. Oxford Reviews of Reproductive Biology, Vol.2. Pp.273. ISBN 0-19-857535-1. (Clarendon Oxford University Press: 1980.) £20.
- FUSTER, J. M. The Prefrontal Cortex: Anatomy, Physiology and Neuropsychology of the Frontal Lobe. Pp.232. ISBN 0-89004-524-0. (Raven: 1980.) \$22.
- HINCHLIFFE, J. R. and JOHNSON, D. R. The Development of the Vertebrate Limb. An Approach through Experiment, Genetics and Evolution. Pp.268. 0-19-857552-1. (Clarendon/Oxford University Press: 1980.) £20.
- JAMES, C. A. Manual of Assessment Keys for Plant Disease Pp.88. Loose leaf. (The American Phytopathological Society, 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) \$10.
- LOCKARD, J. S. and WARD, A. A. (eds). Epilepsy. A Window to Brain Mechanisms. Pp.286. ISBN 0-89004-499-6. (Raven: 1980.) \$29.50.
- MCGANN, B. Behaviour, Health and Lifestyle: The Meaning of Human Health and the Essential Factors Which Enhance It. Pp.349. ISBN 0-906408-14-8. (Villa Books: Dublin, 1980.) £9.50 (UK), £10.45 (Eire).
- MILNE, I. and MILNE, M. The Audubon Society Field Guide to North American Insects and Spiders. Pp. 989. Flexi ISBN 0-394-50763-0. (Knopf: 1980.) \$9.95.
- OTTOWAY, J. H. The Biochemistry of Pollution. The Institute of Biology's Studies in Biology, No. 123. Pp.58. Flexi ISBN 0-7131-2748-8. (Edward Arnold: 1980.) £2.50.
- PANCHEN, A. L. (ed.). The Terrestrial Environment and the Origin of Land Vertebrates. The Systematics Association, Special Volume No.15. Proceedings of an International Symposium held at the University of Newcastle upon Tyne. Pp.634. ISBN 0-12-544780-9. (Academic: 1980.) £28, \$87.50.
- PIOMELLI, S. and YACHNIN, S. (eds). Current Topics in Haematology, Vol.3. Pp.271. ISBN 0-8451-0352-0. (Alan R. Liss: New York, 1980.) Np.
- POLUNIN, O. Flowers of Greece and the Balkans: A Field Guide. Pp.568. ISBN 0-19-217626-9. (Oxford University Press: 1980.) £40.
- PORTER, R. and WHELAN, F. (eds). Metabolic Activities of the Lung. Ciba Foundation Symposium 78. Pp.401. ISBN 90-219-4084-1. (Excerpta Medica: 1980.) \$66.25, Dfl.136.
- RATCLIFFE, D. The Peregrine Falcon. Pp.416. ISBN 0-931130-05-0. (Buteo Books: PO Box 481, Vermillion, SD 57069 USA, 1980.) \$42.50.
- RHOADS, D. C. and LUTZ, R. A. (eds). Skeletal Growth of Aquatic Organisms. Biological Records of Environmental Change: Topics in Geobiology, Vol.1. Pp.750. ISBN 0-306-40259-9. (Plenum: 1980.) \$47.50.
- ROTMAN, A. *et al.* (eds). Platelets: Cellular Response Mechanisms and their Biological Significance. Proceedings of an EMBO Workshop held at The Weizmann Institute of Science, Rehovot, Israel, April 1980. Pp.327. ISBN 0-471-27896-3. (Wiley: 1980.) £17.50.
- SALANKI, J. and TURPAEV, T. M. (eds). Neurotransmitters: Comparative Aspects. Pp.514. ISBN 963-05-2601-8. (Akadémiai Kiadó, Publishing House of the Hungarian Academy of Sciences: H-1361 Budapest, P.O. Box 36, 1980.) \$40.
- SANER, G. Chromium in Nutrition and Disease. Current Topics in Nutrition and Disease, Vol.2. Pp.135. ISBN 0-8451-1601-0. (Alan R. Liss: New York, 1980.) \$16.
- SBARRA, A. J. and STRAUSS, R. R. (eds). The Reticuloendothelial System: A Comprehensive Treatise. Vol.2, Biochemistry and Metabolism. Pp.4342. ISBN 0-306-40292-0. (Plenum: 1980.) \$45.
- SCHAAD, N. W. (ed.). Laboratory Guide for Identification of Plant Pathogenic Bacteria. Pp.76. Loose leaf. ISBN 0-89054-028-4. (The American Phytopathological Society: 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) \$10.
- SIGEL, H. (ed.) Metal Ions in Biological Systems. Vol.11: Metal Complexes as Anticancer Agents. Pp.440. ISBN 0-8247-1004-5. (Marcel Dekker: 1980.) SwFr.115.
- SKULACHEV, V. P. Biology Reviews (Physico-Chemical Aspects), Vol.1 (1980). Soviet Scientific Reviews, Section D. Pp.476. ISBN 3-7186-0020-A. (Harwood Academic Publishers: London, 1980.) Np.
- SMITH, A. D., LLINAS R. and KOSTYUK, P. G. (eds). Commentaries in the Neurosciences. Pp.668. ISBN 0-08-025501-9. (Pergamon: 1980.) £30, \$67.50.
- STACE, C. A. Plant Taxonomy and Biosystematics. Pp.279. Flexi ISBN 0-7131-2802-X. (Edward Arnold: 1980.) £8.95.
- STEIN, R. B. Nerve and Muscle: Membranes, Cells and Systems. Pp.265. ISBN 0-306-40512-1. (Plenum: 1980.) \$18.95.
- TAPPAN, H. The Paleobiology of Plant Protists. Pp.1028. ISBN 0-7167-1109-5. (W. H. Freeman: 1980.) \$95.
- TURNER, N. C. and KRAMER, P. J. (eds). Adaptation of Plants to Water and High Temperature Stress. Pp.482. ISBN 0-471-05372-4. (Wiley: 1980.) £21.50.
- WARNER, N. L. (ed.). Contemporary Topics in Immunobiology, Vol.11. Pp.299. ISBN 0-306-40419-2. (Plenum: 1980.) \$32.50.
- WEETE, J. D. Lipid Biochemistry of Fungi and Other Organisms. Pp.388. ISBN 0-306-40570-9. (Plenum: 1980.) \$45.
- WEINER, H., HOFER, M. A. and STUNKARD, A. J. (eds). Brain, Behaviour, and Bodily Disease. Research Publications: Association for Research in Nervous and Mental Disease, Vol.59. Pp.388. ISBN 0-89004-480-5. (Raven: 1980.) \$37.
- WHATLEY, J. M. and WHATLEY, F. R. Light and Plant Life. The Institute of Biology's Studies in Biology, No. 124. Pp.91. Flexi ISBN 0-7131-2785-6. (Edward Arnold: 1980.) £3.40.
- WHITAKER, J. O. Jr The Audubon Society Field Guide to North American Mammals. Pp.745. Flexi ISBN 0-394-50762-2. (Knopf: 1980.) \$9.95.

- WHITEMAN, P. C. Tropical Pasture Science. Pp.392. Hbk ISBN 0-19-859470-4; pbk ISBN 0-19-859471-2. (Oxford University Press: 1980.) Hbk £20; pbk £5.90.
- WILDENTHAL, K. (ed.). Degradative Processes in Heart and Skeletal Muscle. Research Monographs in Cell and Tissue Physiology, Vol. 3. Pp.456. ISBN 0-444-80235-5. (Elsevier/North-Holland Biomedical: 1980.) \$95, Dfl. 195.
- WINFREE, A. T. The Geometry of Biological Time. Biomathematics, Vol. 8. Pp.530. ISBN 3-540-09373-7. (Springer-Verlag: 1980.) DM 59.50, \$32.80.
- ZENTMYER, G. A. *Phytophthora cinnamomi* and the Diseases It Causes. Monograph No.10. Pp.96. Pbk ISBN 0-89054-030-6. (The American Phytopathological Society: 3340 Pilot Knob Road, St. Paul, Minnesota 55121, 1980.) \$8.

Applied Biological Sciences

- BRIDGES, J. W. and CHASSEAUD, L. F. (eds). Progress in Drug Metabolism, Vol.5. Pp.358. ISBN 0-471-27776-2. (Wiley: 1980.) £28.50.
- COAKER, T. H. (ed.). Applied Biology, Vol. V. Pp.407. ISBN 0-12-040905-4. (Academic: 1980.) £24.50, \$56.50.
- DRAPER, H. H. (ed.). Advances in Nutritional Research, Vol.3. Pp.362. ISBN 0-306-40212-0. (Plenum: 1980.) \$32.50.
- DURIG, J. R. (ed.). Analytical Applications of FT-IR to Molecular and Biological Systems. Proceedings of the NATO ASI held at Florence, Italy, August 1980. NATO Advanced Study Institutes Series C: Mathematical and Physical Sciences 57. Pp.607. ISBN 90-277-1145-3. (Riedel: 1980.) Dfl.130, \$68.50.
- LEHNER, T. and CIMASONI, G. (eds). The Borderland between Caries and Periodontal Disease II. Proceedings of the 2nd European Symposium held in Geneva, Switzerland, February 1980. Pp.288. ISBN 0-12-792506-6. (Academic: 1980.) £18.20, \$42.
- NAIR, S. *et al.* (eds). Computers in Critical Care and Pulmonary Medicine. Proceedings of the 1st International Symposium, held at Norwalk, Connecticut, May 1979. Pp.427. ISBN 0-306-40449-4. (Plenum: 1980.) \$39.50.
- RAPADO, A., WATTS, R. W. E. and DeBRUYN, C. H. M. M. (eds). Purine Metabolism in Man—III: Clinical and Therapeutic Aspects. Advances in Experimental Medicine and Biology, Vol. 122B. Proceedings of the first half of the 3rd International Symposium, held in Madrid, June 1979. Pp.444. ISBN 0-306-40310-2. (Plenum: 1980.) \$45.
- RAUS, J., MARTENS, H. and LECLERCQ, G. (eds). Cytotoxic Estrogens in Hormone Receptive Tumors. Based on the Proceedings of a Workshop held at Diepenbeek, Belgium, September 1979. Pp.285. ISBN 0-12-583050-5. (Academic: 1980.) £14.80, \$34.50.
- SOMOGYI, J. C. and TARJAN, R. (eds). Foreign Substances and Nutrition. Bibliotheca Nutritio et Dieta, No. 29. 16th Symposium of the Group of European Nutritionists, Budapest, November 1978. Pp.132. Flexi ISBN 3-8055-0621-X. (Karger: Basel, 1980.) Sw Fr. 83, DM 99, \$49.75.
- TWYCCROSS, R. G. and VENTAFRIDA, V. (eds). The Continuing Care of Terminal Cancer Patients. Proceedings of an International Seminar, Milan, October, 1979. Pp.280. ISBN 0-80-024943-4. (Pergamon: 1980.) £18, \$40.

Psychology

- ANDERSON, J. R. Cognitive Psychology and its Implications. Pp.503. ISBN 0-7167-1197-4. (W. H. Freeman: 1980.) £8.10.
- MOWRER, O. H. (ed.). Psychology of Language and Learning. Pp.294. ISBN 0-306-40371-4. (Plenum: 1980.) \$27.50.
- RIEVER, R. W. (ed.). Applied Psycholinguistics and Mental Health. Applied Psycholinguistics and Communication Disorders. Pp.224. ISBN 0-306-40392-7. (Plenum: 1980.)

General

- AHRENS, L. H. (ed.). Chemistry of the Moon. Physics and Chemistry of the Earth, Vol. 10. Pp.264. ISBN 0-08-020287-X. (Pergamon: 1980.) £43.
- BAILEY, W. Man, Religion, and Science: A Functional View. Pp.242. Pbk ISBN 0-9604196-0-8. (Wm. Bailey: Santa Barbara, 1980.) \$9.95.
- BETTS, R. J. Business Economics for Engineers. Pp.246. Flexi ISBN 0-07-084112-8. (McGraw-Hill: 1980.) Np.
- BOYLE, R. H. Bass. Pp.160. ISBN 0-393-01379-0. (W. W. Norton: 1980.) \$27.50.
- CALLAHAN, D. and BOK, S. (eds). Ethics Teaching in Higher Education. Pp.315. ISBN 0-306-40522-9. (Plenum: 1980.) \$19.50.
- COREA, G. Need for Change: Towards the New International Order. A Selection from Major Speeches and Reports with an Introduction. Pp.278. ISBN 0-08-02095-0. (Pergamon: 1980.) £12, \$25.
- HARTWIG, S. (ed.). Heavy Gas and Risk Assessment. Proceedings of a Symposium on Heavy Gas, Frankfurt/Main, September 1979. Pp.306. ISBN 90-277-1108-9. (Reidel: 1980.) Dfl. 70, \$37.
- MILUNSKY, A. and ANNAS, G. J. (eds). Genetics and the Law, II. Pp.480. ISBN 0-306-40388-9. (Plenum: 1980.) \$29.50.
- PERROT J. and VANDERMEERSCH, B. (eds). Paléorient, Vol. 5. 1979. Pp.310. Flexi ISBN 2-222-02613-X. (Éditions Du Centre National de la Recherche Scientifique: Paris, 1980.) Np.
- PRASAD, E. A. V. Ground Water in Varahamihira's Brhat Samhita. Pp.351. (E. A. V. Prasad, Sri Venkateswara University, Tirupati, India: 1980.) Rs. 125.
- ROBINS, E. Secret Eden: Africa's Enchanted Wilderness. Pp.128. ISBN 0-241-10423-8. (Elm Tree: London, 1980.) £8.95.
- SCHNEIDER, A. P. and ABROMS, I. F. The Practical Management of the Developmentally Disabled Child. Pp.461. ISBN 0-8016-0061-8. (C. V. Mosby: 1980.) Np.